

PCT/GB01/00299

WO 01/55386

1

Toll-like ReceptorField of the invention

The present invention relates to a novel Toll-like receptor or a variant thereof. A variant may demonstrate Toll-like receptor activity such as activation of NF κ B, or may inhibit Toll-like receptor activity.

Background of the invention

A family of human Toll-like receptors has been described in the literature. These receptors are termed Toll-like receptors in view of common structural features shared with a Drosophila Toll (dToll) receptor molecule which is involved in embryonic development. Toll and Toll-like receptors are type I transmembrane proteins, with extracellular leucine rich repeat motifs and an intracellular signalling domain homologous to that of members of the interleukin 1 receptor superfamily.

Drosophila Toll also plays an important role in the adult fly and is involved in immune surveillance mechanisms required for recognition of bacterial and fungal pathogens and regulation of specific innate immune defence gene expression. Activation of dToll receptors in response to infection by specific micro-organisms is thought to require the production of a protein ligand called Spaetzle. The human Toll-like receptors (hTLRs) are also thought to participate in mechanisms of innate immunity and inflammation acting as pattern recognition receptors (PRRs) for bacteria and other micro-organisms. hTLRs are expressed on antigen presenting cells including monocytes and dendritic cells and like dToll play roles in innate immunity. TLRs can elicit pro-inflammatory cytokine production and induce expression of cell surface co-stimulatory receptors required for activation of T-cells. Some hTLRs may help to co-ordinate interactions between cells of the innate and acquired immune systems to orchestrate an integrated immune response to infection.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 August 2001 (02.08.2001)

(10) International Publication Number
WO 01/55386 A1

PCT

(31) International Patent Classification: C12N 15/12, 15/63, 5/10, G01N 33/68, C07K 14/705, 16/28, A61K 31/7083, 39/17, 48/00

(32) International Filing Date: 25 January 2001 (25.01.2001)

(33) International Filing Date: 25 January 2001 (25.01.2001)

(34) Filing Language: English

(35) Publication Language: English

(36) Priority Date: 23 January 2000 (23.01.2000) GB 0001704.6

(71) Applicant (for all designated States except US): GLAXO-SMITHKLINE (GBGB); Corporate Intellectual Property, Two New Horizons court, Brentford, Middlesex TW8 9BP (GB).

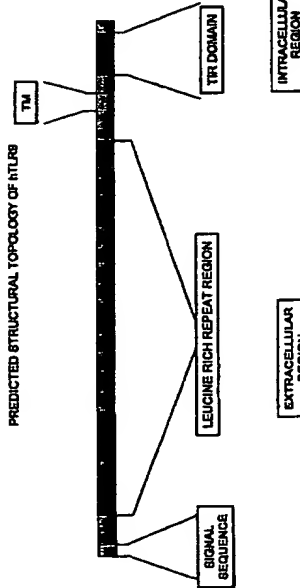
(72) Inventors and
(73) Inventsors/Applicants (for US only): LEWIS, Alan, Peter (GBGB); Glaxo Wellcome plc, Gunpowder Wood Road, Stevenage, Hertfordshire SG1 2NY (GB); BAY, Keith, Paul (GBGB); Glaxo Wellcome plc, Gunpowder Wood Road, Stevenage, Hertfordshire SG1 2NY (GB).

Published:
with international search report
before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guide to the Preparation of the PCT Application" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: TOLL-LIKE RECEPTOR

PREDICTED STRUCTURAL TOPOLOGY OF hTLR



(57) Abstract: An isolated Toll-like receptor polypeptide comprises the amino acid sequence of SEQ ID NO: 2, a variant or a fragment thereof which has immunomodulatory activity. Polynucleotides encoding such a Toll-like receptor are also described. A method for identification of a substance that modulates Toll-like receptor activity comprises contacting a polypeptide of the invention with a test substance and monitoring for immunomodulatory activity.

The specific functions of two mammalian TLRs, TLR2 and TLR4, have recently been identified. TLR2 and TLR4 are involved in mediating host responses to gram positive and gram negative bacteria through recognition of specific bacterial wall components. It has also recently been shown that TLR4 mediates responses to certain viral proteins including respiratory syncytial virus (RSV) (Nature Immunology 1: 398 2000).

Additionally TLRs may form heterodimeric functional complexes. Components of the intracellular signal transduction pathways of some hTLRs appear to be shared with Interleukin-1 (IL-1) receptor transduction pathways. Stimulation of TLR2 and TLR4 leads to activation of NF- κ B via an adapter protein MyD88 and recruitment of the IL-1 receptor associated kinases (IRAKs).

Summary of the invention

A novel Toll-like receptor is now provided which is a screening target for the identification and development of novel pharmaceutical agents which modulate the activity of the receptor and in particular have immunomodulatory activity. These agents may be used in the treatment and/or prophylaxis of inflammatory diseases, cardiovascular diseases, systemic infections and autoimmune diseases, such as asthma, rhinitis, chronic obstructive pulmonary disease (COPD), emphysema, inflammatory bowel disease such as ulcerative colitis and Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, Alzheimer's disease, atherosclerosis, viral, fungal and bacterial infections, septic shock syndrome associated with systemic infection involving gram positive and gram negative bacteria, diabetes, Multiple Sclerosis. These agents may also be used as immunoadjuvants to enhance or alter the immune response in vaccine therapy.

Accordingly, the present invention provides an isolated Toll-like-receptor polypeptide which comprises:

- (i) the amino acid sequence of SEQ ID NO: 2;
- (ii) a variant of (i) which has immunomodulatory activity; or
- (iii) a fragment of (i) or (ii) which retains immunomodulatory activity.

Preferably, a variant has at least 80% identity to the amino acid sequence of SEQ ID NO: 2, more preferably at least 85% identity therewith, for example 87% identity therewith.

The invention also provides a polynucleotide encoding a polypeptide of the invention. Such a polynucleotide may be a polynucleotide which encodes a Toll-like receptor polypeptide which has immunomodulatory activity. The polynucleotides of the invention may be DNA or RNA, for example mRNA. A polynucleotide according to the invention comprises:

- (a) the nucleic acid sequence of SEQ ID NO: 1 and/or a sequence complementary thereto;
- (b) a sequence which hybridises under stringent conditions to a sequence as defined in (a);
- (c) a sequence that is degenerate as a result of the genetic code to a sequence as defined in (a) or (b); or
- (d) a sequence having at least 80% percent identity to a sequence as defined in (a), (b) or (c).

The present invention also provides a polypeptide expressed from a polynucleotide according to (a), (b), (c) or (d) above, in particular a polypeptide comprising a toll-like receptor according to the invention, encoded by the mRNA derived from a DNA sequence according to (a) or (b) above, thus the invention provides an isolated toll-like receptor polypeptide which is obtainable by expression *in vitro* or *in vivo* of a DNA molecule comprising the sequence of nucleotides as shown in SEQ ID NO: 1.

The polypeptides of SEQ ID NO 2 and SEQ ID NO 4, which are different isoforms expressed from the multiple exon *tlr9* gene, are herein referred to as TLR9 and TLR9-A, respectively. TLR9-A is encoded by the nucleotide sequence of SEQ ID No. 3, which is encoded within SEQ ID No.1 except for the Initiating methionine, that is encoded by a second exon as illustrated in figure 1. (see Hemmi et al. Nature 408, 740-745 2000; Du et al. European Cytokine Network, 11: 382-371, 2000; Chuang and Ulevitch, European Cytokine Network, 11: 372-378 for isolation of the cDNAs, and corresponding sequence database accessions EMBL:AB045180, EMBL:AF259282, EMBL:AF245704).

10

In further aspects of the invention we provide:

- an expression vector capable of expressing a polypeptide of the invention comprising a polynucleotide as defined above.
- a host cell comprising an expression vector of the invention.
- an antibody specific for a polypeptide of the invention.
- a method for identification of a compound that modulates Toll-like receptor activity, comprising contacting a polypeptide of the invention with a test compound and monitoring for immunomodulatory activity.

15

20. Compounds which are identifiable in accordance with this method may be used in the treatment of a subject having a disorder that is responsive to Toll-like receptor modulation such as an inflammatory or cardiovascular disorder or systemic infection or autoimmune disease, including asthma, chronic obstructive pulmonary disease (COPD), emphysema, inflammatory bowel disease such as ulcerative colitis and Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, viral, fungal and bacterial infections, Alzheimer's disease, atherosclerosis, septic shock syndrome associated with systemic infection involving gram positive and gram negative bacteria, diabetes and Multiple Sclerosis. In particular, compounds which are identifiable in accordance with this method may be used in the treatment of a subject having allergic asthma or

30

rinitis. Further, such compounds may have immunomodulatory activity and be of use in the treatment of, or as adjuvants in vaccination against, bacterial or viral infections or as components of anti-cancer vaccines.

5 Compounds identifiable in accordance with this method include, in particular, synthetic or naturally occurring oligopeptides or polypeptides, oligonucleotides or polynucleotides which bind directly to the Toll-like receptor of the invention, and synthetic or naturally occurring oligopeptides or polypeptides, oligonucleotides or polynucleotides which modulate the Toll-like receptor of the present invention via one or more intermediate signal transducers. Such oligo- or polynucleotides may be "CG-rich" sequences or sequences including one or more unmethylated CpG nucleotide pairs.

10

15 In an alternative aspect of the invention, a polypeptide comprises a fragment or variant of SEQ ID NO 2 which is capable of inhibiting the activity of TLR9 or TLR9-A, for use in the treatment of an immune or inflammatory disorder.

In a further aspect of the invention, a polypeptide or polynucleotide in accordance with the invention or a compound identifiable in accordance with the invention is provided for use as an adjuvant or as an immunotherapeutic agent, for example in a vaccine.

20

Brief Description of the Sequences

SEQ ID NO: 1 is the amino acid sequence of human protein TLR9 and its encoding DNA;

25

SEQ ID NO: 2 is the amino acid sequence alone of TLR9;

SEQ ID NO: 3 is the amino acid sequence of human protein TLR9-A and its encoding cDNA (EMBL:AF259282);

SEQ ID NO: 4 is the amino acid sequence alone of TLR9-A (Hemmi et al.).

30

Brief Description of the Drawings

Fig. 1 is a diagrammatic illustration showing the exon arrangement encoding TLR9 and TLR9-A;

Fig. 2 shows tissue distribution data for TLR9 (using a human tissue cDNA plate). The profile shows predominant expression in tonsil and adenoid tissues with lower levels of expression detected in adipose, adrenal, foetal brain, cerebellum, jejunum, lung, myometrium, omentum, head of pancreas, rectum, skeletal muscle, spleen and thymus tissues;

Fig. 3 shows tissue distribution data for TLR9 (using a human disease cDNA plate). The profile shows predominant expression in lung tissue, bone marrow and PBMC with lower levels of expression detected in some colon, breast and brain/cerebellum samples;

Fig. 4 illustrates, in diagrammatic form, the predicted structural topology of human TLR9 - "TM" is the transmembrane portion, "TIR" is the cytosolic region conserved among interleukin and toll-like receptors known as the Toll Interleukin Receptor domain.

Detailed Description of the Invention

A single open reading frame was identified in genomic DNA (SEQ ID No.1), which encodes a protein of 1055 residues, predicted molecular weight 118,515 (PeptideSort - GCG Software) and the amino acid sequence shown in SEQ ID No.2. This sequence included the TIR domain common to Toll-like receptors and members of the interleukin-1 receptor family e.g. IL1RI, and the N-terminal sequence contains structural features as shown in figure 4. These features include, in order from the N- to the C-terminus, a predicted signal sequence with a potential cleavage site between residues 48 and 49 or 50 and 51 (SPScan in GCG; SignalP), a leucine-rich repeat motif domain, a potential transmembrane region and the TollIL-1R homologous region (TIR; Pfam Database). These motifs confirmed that the protein was likely to be expressed as a single transmembrane receptor-like molecule belonging to the TLR rather than the IL1R

family and therefore, it was designated TLR9 based on existing published and in-house nomenclature.

Throughout the present specification and the accompanying claims the words "comprise" and "include" and variations such as "comprises", "comprising", "includes" and "including" are to be interpreted inclusively. That is, these words are intended to convey the possible inclusion of other elements or integers not specifically recited, where the context allows.

The present invention relates to a human Toll-like receptor, referred to herein as TLR9, and variants or fragments thereof. Sequence information for TLR9 is provided in SEQ ID NO: 1 (nucleotide and amino acid) and in SEQ ID NO: 2 (amino acid only). A polypeptide of the invention consists essentially of the amino acid sequence of SEQ ID NO: 2 or of a functional variant of that sequence. One important variant of TLR9 is TLR9-A, sequence information for which is provided in SEQ ID NO: 3 (nucleotide and amino acid) and in SEQ ID NO: 4 (amino acid only).

The polypeptides are provided in isolated form. The term "isolated" is intended to convey that the polypeptide is not in its native state, insofar as it has been purified at least to some extent or has been synthetically produced, for example by recombinant methods. The polypeptide may be mixed with carriers or diluents which will not interfere with the intended purpose of the polypeptide and still be regarded as substantially isolated. The term "isolated" therefore includes the possibility of the polypeptide being in combination with other biological or non-biological material, such as cells, suspensions of cells or cell fragments, proteins, peptides, expression vectors, organic or inorganic solvents, or other materials where appropriate, but excludes the situation where the polypeptide is in a state as found in nature.

A polypeptide of the invention may also be in a substantially purified form, in which case it will generally comprise the polypeptide in a preparation in which more than 50%, e.g. more than 80%, 90%, 95% or 99%, by weight of the polypeptide in the preparation is a polypeptide of the invention. Routine methods, can be employed to purify and/or synthesise the proteins according to the invention. Such methods are well understood by persons skilled in the art, and include techniques such as those disclosed in Sambrook *et al.*, Molecular Cloning: a Laboratory Manual, 2nd Edition, CSH Laboratory Press (1989), the disclosure of which is included herein in its entirety by way of reference.

10

The term "variants" refers to polypeptides which have the same essential character or basic biological functionality as TLR9. The essential character of TLR9 can be defined as that of a Toll-like receptor. In particular, it refers to a polypeptide which has an immunomodulatory activity. In one aspect of the invention, a polypeptide of the invention may activate NF κ B or may elicit pro-inflammatory cytokine production or induce expression of cell surface co-stimulatory receptors required for activation of T-cells.

15

Alternatively, a variant of the polypeptide of the invention is one which exhibits binding to the same ligand as TLR9. Such ligand binding may be assayed using the assays described below.

20

In other aspects of the invention a variant is one which does not show the same function as TLR9 but which may be used to inhibit this function. For example, a variant polypeptide for use in an assay or therapy is one which inhibits TLR9 activity, for example by inhibiting or competing out ligand binding or receptor complex formation by TLR9. Alternatively, a variant may be one which inhibits ligand binding to TLR9. Such a variant may inhibit activation of NF κ B or inhibit cytokine production and expression of cell surface co-stimulatory receptors.

25

Such inhibitors may be used as immunomodulators to reduce inappropriate TLR activation in asthma or other chronic inflammatory diseases, or septic shock.

5

To determine whether a variant has the same essential function as TLR9, the immunomodulatory activity can be determined by monitoring the effect of a substance on different immune responses. For example the effect of the substance under test on NF κ B activation mediated through binding the polypeptide of the present invention may be monitored. This can be carried out, for example, by co-transfection of a construct expressing the polypeptide with a construct containing a reporter gene, such as secreted placental alkaline phosphatase, under the control of a suitable NF κ B-responsive promoter and monitoring for expression of the reporter gene.

10

Alternatively, other immunomodulatory activity such as the production of cytokines can be determined by monitoring cytokine production following incubation of a test substance with a cell expressing a polypeptide of the invention. Such assays may be carried out in the presence or absence of additional T-lymphocytes to assess the effect of such cytokines, or the direct action of a polypeptide of the invention, on such T-lymphocytes to thus determine immunomodulatory activity.

15

20

Alternatively, the Toll-like receptor functionality is as a peptide which binds a ligand of TLR9, inhibits immunomodulatory activity by TLR9 or inhibits ligand binding to TLR9 and can be determined by an assay as described below.

25

Preferably, a polypeptide of the invention will show the structural features associated with a Toll-like receptor. Preferably, a polypeptide of the invention, or a functional fragment thereof, contains one or more of the following structural features associated with a Toll-like receptor: an extracellular region containing leucine-rich repeat motif and cysteine-rich regions involved in ligand binding; a

30

10

single hydrophobic transmembrane region; and a C-terminal cytoplasmic domain sharing homology with other TLRs and members of the IL-1 receptor family.

Typically, polypeptides with more than about 65% identity, preferably at least 80% or at least 90% and particularly preferably at least 95%, at least 97%, or at least 99% identity, with the amino acid sequences of SEQ ID NO: 2 over a region of at least 20, preferably at least 30, at least 40, at least 60 or at least 100 contiguous amino acids or over the full length of SEQ ID NO: 2, are considered as variants of the proteins. Identity is calculated using the widely used GCG (University of Wisconsin) suite of programs and preferably using the distances software (correction method). Such variants may include allelic variants and the deletion, modification or addition of single amino acids or groups of amino acids within the protein sequence, as long as the peptide maintains the basic biological functionality of the Toll-like receptor, having a similar function to TLR9 or inhibits such function such as preventing ligand binding or TLR9 mediated activation. Such variants also include isoforms such as TLR9-A, which is 23 amino acids (or 2.2%) shorter than TLR9 (see SEQ ID NO: 2 and SEQ ID NO: 4) and thus shows 97.8 identity therewith. Transcription of the nucleotide sequence presented in SEQ ID NO: 1 can result, due to variable mRNA splicing involving a second exon encoding an alternative initiating methionine, in an mRNA having the sequence of SEQ ID NO: 3 which, when translated, results in the polypeptide TLR9-A depicted in SEQ ID NO: 4.

Amino acid substitutions may be made, for example from 1, 2 or 3 to 10, 20 or 30 substitutions. The modified polypeptide generally retains activity as a TLR9 receptor or inhibitor of TLR9 receptor activity. Conservative substitutions may be made, for example according to the following Table. Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other.

11

ALIPHATIC	Non-polar	G A P
		I L V
	Polar-uncharged	C S T M
	Polar-charged	N Q D E K R H F W Y
AROMATIC		

Shorter polypeptide sequences are within the scope of the invention. For example, a peptide of at least 20 amino acids or up to 50, 60, 70, 80, 100 or 150 amino acids in length is considered to fall within the scope of the invention as long as it demonstrates the basic biological functionality of TLR9 or inhibits TLR9. In accordance with this aspect of the invention the peptide may also comprise an epitope of TLR9 for generation of antibodies. In particular, but not exclusively, this aspect of the invention encompasses the situation when the protein is a fragment of the complete protein sequence and may represent a ligand-binding region (N-terminal extracellular domain) or an effector binding region (C-terminal intracellular domain). Fragments from which the C-terminus has been removed may be used as decoy receptors. Other fragments such as a secreted or soluble form of the receptor may be generated for use in an assay or in therapy in accordance with the invention. Such fragments can also be used to raise anti-TLR9 antibodies.

Polypeptides of the invention may be chemically modified, e.g. post-translationally modified. For example, they may be glycosylated or may comprise modified amino acid residues. They may also be modified by the addition of histidine residues or an epitope tag for example by a (His)₆ or (His)₆ sequence or a HA, TT, Myc or Flag tag to assist their purification or detection. They may be modified by the addition of a signal sequence to promote insertion into the

cell membrane. Such modified polypeptides fall within the scope of the term "polypeptide" of the invention.

6 The invention also includes nucleotide sequences that encode for TLR9 or variants thereof as well as nucleotide sequences which are complementary thereto. The nucleotide sequence may be RNA or DNA including genomic DNA, synthetic DNA or cDNA. Preferably the nucleotide sequence is a DNA sequence and most preferably, a cDNA sequence. Nucleotide sequence information is provided in SEQ ID NO: 1. Such nucleotides can be isolated from human cells or synthesised according to methods well known in the art, as described by way of example in Sambrook *et al.* Such nucleotides can typically be isolated from activated cells of the immune system, heart, lung, pancreatic islet cells and lymph nodes, adenoid and tonsil tissues. Figures 2 and 3 show the tissue distribution of RNA encoding TLR9, as determined by extraction of total RNA from normal or disease tissue or cells which is then used to generate cDNA for real time quantitative PCR using suitable primers and probes (TaqMan analysis) to assess expression patterns. The profiles show differential expression across tissues tested and predominance to sites containing inflammatory cell types.

20 Typically a polynucleotide of the invention comprises a contiguous sequence of nucleotides which is capable of hybridising under selective conditions to the coding sequence or the complement of the coding sequence of SEQ ID NO: 1.

25 A polynucleotide of the invention can hybridize to the coding sequence or the complement of the coding sequence of SEQ ID NO: 1 (or of SEQ ID NO: 3) at a level significantly above background. Background hybridisation may occur, for example, because of other cDNAs present in a cDNA library. The signal level generated by the interaction between a polynucleotide of the invention and the coding sequence or complement of the coding sequence of SEQ ID NO: 1 or of SEQ ID NO: 3 is typically at least 10 fold, preferably at least 100 fold, as intense

as interactions between other polynucleotides and the coding sequence of SEQ ID NO: 1 or of SEQ ID NO: 3. The intensity of interaction may be measured, for example, by radiolabelling the probe, e.g. with ^{32}P . Selective hybridisation may typically be achieved using conditions of low stringency (0.03M sodium chloride and 0.03M sodium citrate at about 40°C), medium stringency (for example, 0.03M sodium chloride and 0.03M sodium citrate at about 50°C) or high stringency (for example, 0.03M sodium chloride and 0.03M sodium citrate at about 60°C).

10 The coding sequence of SEQ ID NO: 1 may be modified by nucleotide substitutions, for example from 1, 2 or 3 to 10, 25, 50 or 100 substitutions. The polynucleotides of the present invention may alternatively or additionally be modified by one or more insertions and/or deletions and/or by an extension at either or both ends. The modified polynucleotide generally encodes a polypeptide which has Toll-like receptor activity or inhibits the activity of TLR9. Degenerate substitutions may be made and/or substitutions may be made which would result in a conservative amino acid substitution when the modified sequence is translated, for example as shown in the Table above.

20 A nucleotide sequence of the invention which is capable of selectively hybridising to the complement of the DNA coding sequence of SEQ ID NO: 1 will generally have at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98% or at least 99% sequence identity to the coding sequence of SEQ ID NO: 1 over a region of at least 20, preferably at least 30, for instance at least 40, at least 60, more preferably at least 100 contiguous nucleotides or most preferably over the full length of SEQ ID NO: 1. Methods of measuring nucleic acid and protein homology are well known in the art. For example the UWGCG Package provides the BESTFIT program which can be used to calculate homology (Devereux *et al* 1984). Similarly the PILEUP and BLAST algorithms can be used to line up sequences (for example are described in Altschul 1993,

and Altshul *et al* 1990). Many different settings are possible for such programs. In accordance with the invention, the default settings may be used.

Any combination of the above mentioned degrees of sequence identity and minimum sizes may be used to define polynucleotides of the invention, with the more stringent combinations (i.e. higher sequence identity over longer lengths) being preferred. Thus, for example a polynucleotide which has at least 80% sequence identity over 25, preferably over 30 nucleotides forms one aspect of the invention, as does a polynucleotide which has at least 85% sequence identity over 40 nucleotides. The most preferred sequences have at least 70% sequence identity over at least 70% of the full length of the sequence provided by SEQ ID NO: 1.

The nucleotides according to the invention have utility in production of the proteins according to the invention, which may take place *in vitro*, *in vivo* or *ex vivo*. The nucleotides may be involved in recombinant protein synthesis or indeed as therapeutic agents in their own right, utilised in gene therapy techniques. Nucleotides complementary to those encoding TLR9, or antisense sequences, may also be used in gene therapy, such as in strategies for down regulation of expression of the proteins of the invention.

Polynucleotides of the invention may be used as a primer, e.g. a PCR primer, a primer for an alternative amplification reaction, a probe e.g. labelled with a revealing label by conventional means using radioactive or non-radioactive labels, or the polynucleotides may be cloned into vectors.

Such primers, probes and other fragments will preferably be at least 10, preferably at least 15 or at least 20, for example at least 25, at least 30 or at least 40 nucleotides in length. They will typically be up to 40, 50, 60, 70, 100 or 150 nucleotides in length. Probes and fragments can be longer than 150

nucleotides in length, for example up to 200, 300, 400, 500 nucleotides in length, or even up to a few nucleotides, such as five or ten nucleotides, short of the coding sequence of SEQ ID NO: 1.

The present invention also includes expression vectors that comprise nucleotide sequences encoding the proteins or variants thereof of the invention. Such expression vectors are routinely constructed in the art of molecular biology and may for example involve the use of plasmid DNA and appropriate initiators, promoters, enhancers and other elements, such as for example polyadenylation signals which may be necessary, and which are positioned in the correct orientation, in order to allow for protein expression. Other suitable vectors would be apparent to a person skilled in the art. By way of further example in this regard we refer to Sambrook *et al*.

Polynucleotides according to the invention may also be inserted into the vectors described above in an antisense orientation in order to provide for the production of antisense RNA. Antisense RNA or other antisense polynucleotides may also be produced by synthetic means. Such antisense polynucleotides may be used as test compounds in the assays of the invention or may be useful in a method of treatment of the human or animal body by therapy.

Preferably, a polynucleotide of the invention or for use in the invention in a vector is operably linked to a control sequence which is capable of providing for the expression of the coding sequence by the host cell, i.e. the vector is an expression vector. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A regulatory sequence, such as a promoter, "operably linked" to a coding sequence is positioned in such a way that expression of the coding sequence is achieved under conditions compatible with the regulatory sequence.

The vectors may be for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of the said polynucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial plasmid or a resistance gene for a fungal vector. Vectors may be used *in vitro*, for example for the production of DNA or RNA or used to transfect or transform a host cell, for example, a mammalian host cell. The vectors may also be adapted to be used *in vivo*, for example in a method of gene therapy.

Promoters and other expression regulation signals may be selected to be compatible with the host cell for which expression is designed. For example, yeast promoters include *S. cerevisiae* GAL4 and ADH promoters, *S. pombe* *nmt1* and *adh* promoter. Mammalian promoters include the metallothionein promoter which can be induced in response to heavy metals such as cadmium. Viral promoters such as the SV40 large T antigen promoter or adenovirus promoters may also be used. All these promoters are readily available in the art.

Mammalian promoters, such as β -actin promoters, may be used. Tissue-specific promoters may be used. Viral promoters may also be used, for example the Moloney murine leukaemia virus long terminal repeat (MMLV LTR), the rous sarcoma virus (RSV) LTR promoter, the SV40 promoter, the human cytomegalovirus (CMV) IE promoter, adenovirus, HSV promoters (such as the HSV IE promoters), or HPV promoters, particularly the HPV upstream regulatory region (URR). Viral promoters are readily available in the art.

The vector may further include sequences flanking the polynucleotide which comprise sequences homologous to eukaryotic genomic sequences, preferably mammalian genomic sequences, or viral genomic sequences. This will allow the

Introduction of the polynucleotides of the invention into the genome of eukaryotic cells or viruses by homologous recombination. In particular, a plasmid vector comprising the expression cassette flanked by viral sequences can be used to prepare a viral vector suitable for delivering the polynucleotides of the invention to a mammalian cell. Other examples of suitable viral vectors include herpes simplex viral vectors and retroviruses, including lentiviruses, adenoviruses, adeno-associated viruses and HPV viruses (such as HPV-16 or HPV-18). Gene transfer techniques using these viruses are known to those skilled in the art. Retrovirus vectors for example may be used to stably integrate the polynucleotide giving rise to the RNA into the host genome. Replication-defective adenovirus vectors by contrast remain episomal and therefore allow transient expression.

The invention also includes cells that have been modified to express the TLR9 polypeptide or a variant thereof. Such cells include transient, or preferably stable higher eukaryotic cell lines, such as mammalian cells or insect cells, lower eukaryotic cells, such as yeast or prokaryotic cells such as bacterial cells. Particular examples of cells which may be modified by insertion of vectors encoding for a polypeptide according to the invention include mammalian HEK293T, CHO, HeLa and COS cells. Preferably the cell line selected will be one which is not only stable, but also allows for mature glycosylation and cell surface expression of a polypeptide. Cells such as T-cells, monocytes or dendritic cells expressing the receptor may be used for example in screening. Expression may be achieved in transformed oocytes. A polypeptide of the invention may be expressed in cells of a transgenic non-human animal, preferably a mouse. A transgenic non-human animal expressing a polypeptide of the invention is included within the scope of the invention.

It is also possible for the proteins of the invention to be transiently expressed in a cell line or on a membrane, such as for example in a baculovirus expression

18

system. Such systems, which are adapted to express the proteins according to the invention, are also included within the scope of the present invention.

5 According to another aspect, the present invention also relates to antibodies (either polyclonal or preferably monoclonal antibodies, chimeric, single chain, Fab fragments) which have been raised by standard techniques and are specific for a polypeptide of the invention. Such antibodies could for example, be useful in purification, isolation or screening methods involving immunoprecipitation techniques and may be used as tools to further elucidate the function of TLR9 or a variant thereof, or indeed as therapeutic agents in their own right. Antibodies may also be raised against specific epitopes of the proteins according to the invention. Such antibodies may be used to block ligand binding to the receptor. Alternatively an antibody may be provided which acts as an agonist, to cross link receptors of the invention to mediate receptor activity. An antibody, or other compound, "specifically binds" to a protein when it binds with high affinity to the protein for which it is specific but does not bind or binds with only low affinity to other proteins. A variety of protocols for competitive binding or immunoradiometric assays to determine the specific binding capability of an antibody are well known in the art (see for example Maddox *et al* 1993). Such immunoassays typically involve the formation of complexes between the "specific protein" and its antibody and the measurement of complex formation.

25 An important aspect of the present invention is the use of polypeptides according to the invention in screening methods to identify compounds that may act as modulators of Toll-like receptor activity. Any suitable form may be used for the assay to identify a modulator of TLR9 activity. In general terms, such screening methods may involve contacting a polypeptide of the invention with a test compound and then measuring receptor activity.

19

Screening methods may alternatively involve contacting a polypeptide of the invention with a test compound and then monitoring for the effect on immunomodulatory activity.

5 The binding of the substance to a polypeptide in the invention can be determined directly. For example, a radiolabelled test substance can be incubated with a polypeptide of the invention and so that binding of the test substance to the polypeptide can be monitored. Typically, the radiolabelled test substance can be incubated with cell membranes containing the polypeptide until equilibrium is reached. The membranes can then be separated from a non-bound test substance and dissolved in scintillation fluid to allow the radioactive content to be determined by scintillation counting. Non-specific binding of the test substance may also be determined by repeating the experiments in the presence of a saturating concentration of a non-radioactive ligand. Preferably, a binding curve is constructed by repeating the experiment with various concentrations of the test substance.

15 Cell based assays may also be carried out, for example using a cell expressing the TLR9 receptor, and contacting the cell with another cell to look for ligand binding or activation of TLR9-mediated pathways such as NF κ B activation. Alternatively cells expressing TLR9 constitutively may be provided for use in assays for TLR9 function. Such constitutively expressed TLR9 may demonstrate TLR9 activity in the absence of ligand binding. Additional test substances may be introduced in any assay to look for inhibitors of ligand binding or inhibitors of TLR9-mediated activity. Assays are preferably carried out using cells expressing TLR9, and incubating such cells with the test substance optionally in the presence of TLR9 ligand. Alternatively an antibody may be used to complex TLR9 and thus mediate TLR9-activity. Test substances may then be added to assess the effect on such activity.

20

In preferred aspects, a host cell is provided expressing the receptor, or a receptor complex of TLR9 (or TLR9-A) comprising a homodimer, a heterodimer of TLR8 (or TLR9-A) with another Toll-like receptor, or a complex of TLR9 (or TLR9-A) with protein cofactors, and containing an NFkB responsive reporter construct. The host cell is treated with a substance under test for a defined time. The expression of the reporter gene, such as secreted placental (SP) alkaline phosphatase or luciferase is assayed. The assay enables determination of whether the addition of compounds inhibits the induction of the response in target cells. Alternatively the assay may be carried out to identify cytokine production or it may be carried out in the presence of T-cells to identify induction of co-stimulatory receptors required for activation of T-cells.

Assays may also be carried out to identify modulators of receptor shedding. A polypeptide of the invention can be cleaved from the cell surface. Shedding the receptor would act to down regulate receptor signalling. Thus, cell based assays may be used to screen for compounds which promote or inhibit receptor-shedding. Assays may also be carried out to identify substances which modify TLR8 receptor expression for example substances which down regulate expression. Such assays may be carried out for example by using antibodies for TLR8 to monitor levels of TLR8 expression.

Additional control experiments may be carried out. Assays may also be carried out using known ligands of other Toll-like receptors to identify ligands which are specific for polypeptides of the invention. Preferably, the assays of the invention are carried out under conditions which would result in immunomodulatory NFkB mediated activity in the absence of the test substance, to identify inhibitors of Toll-like receptor mediated activity, or agents which inhibit ligand-induced Toll-like receptor activity.

21

Suitable test substances which can be tested in the above assays include combinatorial libraries, defined chemical entities, peptides and peptide mimetics, oligonucleotides and natural product libraries, such as display (e.g. phage display libraries) and antibody products. In a preferred embodiment, the test substance is a variant peptide of the invention. In particularly preferred embodiments, suitable test substances which may be candidate ligands for binding to and modulation of TLR9 or TLR9-A include synthetic or naturally occurring oligonucleotides or polynucleotides which bind directly to the Toll-like receptor or which modulate the Toll-like receptor of the present invention via one or more intermediate signal transducers. Such oligo- or polynucleotides may be "CG-rich" sequences or sequences including one or more unmethylated CpG nucleotide pairs.

The assay may be carried out using full length TLR9 to identify a variant peptide which interferes with TLR9 mediated activity, for example by inhibiting ligand binding.

Test substances may be used in an initial screen of, for example, 10 substances per reaction, and the substances of these batches which show inhibition or activation tested individually. Test substances may be used at a concentration of from 1nM to 1000µM, preferably from 1µM to 100µM, more preferably from 1µM to 10µM.

A protein-binding assay may be developed using a polypeptide of the invention, preferably one encoding the extracellular ligand-binding domain, to identify novel protein ligands of TLR8. Particular examples may be screening of a human cDNA expression library for protein ligands of TLR9 by yeast 2-hybrid protein interaction assay (e.g. as described in International Patent Application No. WO89/49284).

30

22

Another aspect of the present invention is the use of polynucleotides encoding the TLR9 polypeptides of the invention to identify mutations in TLR9 genes which may be implicated in human disorders or to identify cells in which TLR9 is expressed. Identification of such mutations may be used to assist in diagnosis of immune system, lung, kidney, heart or other disorders or susceptibility to such disorders and in assessing the physiology of such disorders. In particular the polynucleotides of the invention may assist in diagnosis of asthma and rheumatoid arthritis. For example, a SNP (single nucleotide polymorphism) has been identified in the genomic DNA encoding TLR9 (G/A nucleotide: The SNP Consortium database accession number TSC0164834). This single base pair change lies in the DNA region encoding the 23 N-terminal residues of TLR9, and this region is spliced out of the mRNA encoding TLR9-A. The nucleotide at this SNP position may affect the efficiency of mRNA splicing in the two different variants - a G at this position may possibly disrupt the splicing machinery and an A might lead to more efficient splicing. Additionally, the presence of a G as compared to an A in an unspliced mRNA would introduce a stop codon and result in different N terminal protein sequences upon translation of that mRNA, thus the two polymorphic variants of the *tlr9* gene may encode receptors which have differing expression levels and/or differing functional activity levels. The present invention provides a diagnostic tool for determining the polymorphic variant in an individual by detecting the DNA sequence at the SNP site. Such a tool may incorporate a nucleotide probe specific for one or other of the polymorphic variants, for example an oligonucleotide of from 5 to 50, preferably 5-20 nucleotides, complementary to a fragment of the nucleotide sequence of SEQ ID No. 1 which extends over the SNP site or a fragment complementary to that sequence with the exception of the single nucleotide change (G to A) at the SNP site. The present invention also provides a method of detecting a polymorphic variant in the *tlr9* gene by determining the sequence of nucleotides at and around the SNP site identified by the SNP consortium database accession number

23

TSC0164834, in particular by determining whether the nucleotide at that SNP site is a G or an A.

Another aspect of the present invention is the use of the compounds that have been identified by screening techniques referred to above in the treatment or prophylaxis of disorders which are responsive to regulation of TLR9 receptor activity. In addition, variant peptides of the invention which inhibit TLR9-mediated activity, for example which inhibit ligand binding or prevent hTLR9 immunomodulatory activity may be used in the treatment or prophylaxis of such disorders. Antibodies which recognise TLR9 may similarly be used in therapy.

In particular, such compounds may be used in the treatment of inflammatory, cardiovascular, systemic infection or autoimmune disease. The compounds may be used to treat bacterial, viral or fungal infections, asthma, rhinitis, chronic obstructive pulmonary disease (COPD), emphysema, an inflammatory bowel disease such as ulcerative colitis and Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, Alzheimer's disease, atherosclerosis, septic shock syndrome associated with systemic infection involving gram positive and gram negative bacteria, diabetes, Multiple Sclerosis.

In an alternative aspect, the invention provides agents which activate TLR9 mediated immunomodulation for use as an immunoadjuvant, or TLR9, and variants thereof, or polynucleotides or DNA encoding a polypeptide of the invention which may be administered for use as immunoadjuvants to enhance or alter the immune response in an individual to an antigen.

The compounds identified according to the screening methods outlined above may be formulated with standard pharmaceutically acceptable carriers and/or excipients as is routine in the pharmaceutical art, and as fully described in Remington's Pharmaceutical Sciences, Mack Publishing Company, Eastern

Pennsylvania 17th Ed. 1985, the disclosure of which is included herein of its entirety by way of reference.

5 The compounds may be administered by enteral or parenteral routes such as via oral, buccal, anal, pulmonary, intravenous, intra-arterial, intramuscular, intraperitoneal, topical or other appropriate administration routes. The dose of a compound to be administered may be determined according to various parameters, especially according to the substance used; the age, weight and condition of the patient to be treated; the route of administration; and the required regimen. Again, a physician will be able to determine the required route of administration and dosage for any particular patient. A typical daily dose is from about 0.1 to 50 mg per kg of body weight, according to the activity of the compound, the age, weight and conditions of the subject to be treated, the type and severity of the disease and the frequency and route of administration.

15 Preferably, daily dosage levels are from 5 mg to 2 g.

Nucleic acid encoding an inhibitor of TLR9 activity may be administered to the mammal. In an alternative aspect of the invention, nucleic acid encoding TLR9 or a variant thereof may be administered to the animal. Such a variant shows immunomodulatory activity of TLR9 such as inducing cytokine production and expression of cell surface co-stimulatory receptors required for activation of T-cells. Nucleic acid, such as RNA or DNA, and preferably, DNA, is provided in the form of a vector, such as the polynucleotides described above, which may be expressed in the cells of the mammal.

25 Nucleic acid encoding the peptide may be administered to the animal by any available technique. For example, the nucleic acid may be introduced by injection, preferably intradermally, subcutaneously or intramuscularly. Alternatively, the nucleic acid may be delivered directly across the skin using a nucleic acid delivery device such as particle-mediated gene delivery. The

nucleic acid may be administered topically to the skin, or to the mucosal surfaces for example by intranasal, oral, intravaginal, intrarectal administration.

5 Uptake of nucleic acid constructs may be enhanced by several known transfection techniques, for example those including the use of transfection agents. Examples of these agents includes cationic agents, for example, calcium phosphate and DEAE-Dextran and lipofectants, for example, lipofectam and transfectam. The dosage of the nucleic acid to be administered can be altered. Typically the nucleic acid is administered in the range of 1pg to 1mg, preferably to 1pg to 10µg nucleic acid for particle mediated gene delivery and 10µg to 1mg for other routes.

Examples

15 Example 1 - Screening for compounds which exhibit protein modulating activity

Mammalian cells, such as HeK293, CHO and HeLa cells over-expressing TLR9 protein are generated for use in the assay. 96 and 384 well plate, high throughput screens (HTS) are employed. TLR9 induced cell activation can be monitored through activation of a signal transduction event (typically resulting in activation of NFκB or AP-1) or transcriptional activation of a reporter gene (typically regulated via NFκB or AP-1 responsive elements). TLR9 induced activation of a reporter gene such as alkaline phosphatase or luciferase can be easily assessed using a suitable colourimetric or fluorimetric assay to measuring production. Such assays may be used to identify receptor antagonists capable of blocking TLR9 activation, inhibitors capable of blocking TLR9 intracellular signal transduction or receptor agonists capable of activating TLR9. Secondary screening involves evaluation of TLR9 modulators in disease related

26

cells. Tertiary screens involve the study of modulators in rat and mouse models of disease relevant to the target.

27

SEQUENCE LISTING

5 <11> GLAXO GROUP LIMITED
<12> PROTEIN
<13> P78721 GCM SER
10 <140>
<141>
<16> 2
<17> PatentIn Ver. 2.1
15 <210> 1
<211> 3165
<212> DNA
<213> Homo sapiens
20 <220>
<221> CDS
<222> (1)..(3165)
25 <400> 1
atg ccc atg aag tgg agt ggg tgg agg tgg acc tgg ggg ccg gcc act 48
Met Pro Met Lys Trp Ser Gly Trp Arg Trp Ser Gly Pro Ala Thr 10
1 5 10 15
cac aca gcc ctc cca ccc cca cag ggt ttc tgc cgc agc gcc ctg cac 96
His Thr Ala Leu Pro Pro Gln Gly Phe Cys Arg Ser Ala Leu His 20
30 25 30
ccg ctg tct ctc ctg gta cag gcc atc atg ctg gcc atg acc ctg gcc 144
Pro Leu Ser Leu Leu Val Gln Ala Ile Met Leu Ala Met Thr Leu Ala 35
40 45
35 ctg gtt acc ttg cct gcc ttc cta ccc tgc ggg ctc cag ccc cac gcc 192
Leu Gly Thr Leu Pro Ala Phe Leu Leu Pro Cys Gln Leu Gln Pro His Gly 50
55 60
ctg gta aac tgc aac tgg ctg ttc ctg aag tct gta ccc cac ttc tcc 240
Leu Val Asn Cys Asn Trp Leu Phe Leu Lys Ser Val Pro His Phe Ser 65
70 75 80
40 atg gca gca ccc ggt gcc aat gtc acc ggc ctt tcc ttg tcc tcc acc 288
Met Ala Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn 85
90 95
45 ggc atc cac cac ctc cat gat tct gac ttt gcc cac ctg ccc agc ctg 336
Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu 100
105 110
cgg cat ctc aac ctc aag tgg aac tgc ccg ccg gtt gcc ctg agc ccc 384
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Pro Val Gly Leu Ser Pro 115
120 125
50 atg cac ttc ccc tgc cac atg acc atc ggg ccc agc acc ttc ttg gct 432
Met His Phe Pro Cys His Met Thr Ile Gln Pro Ser Thr Phe Leu Ala 130
135 140

gag ccc acc ctg gaa gag cta aac cag acc tac aac acc atc atg act 480
Val Pro Thr Leu Glu Glu Leu Ser Tyr Asn Asn Ile Met Thr 155 160
145 gag cct gag ctg ccc aaa tcc cta ttc ctg tcc ctc acc ccc acc 528
5 Val Pro Ala Leu Pro Lys Ser Leu Ile Ser Leu Ser Leu Ser His Thr 175
165 gag cct gag ctg cta gag cct tct gcc acc ctc gcc ctg ctc acc ctg 576
Asn Ile Met Leu Asp Ser Leu Ala Glu Leu His Ala Leu 185 190
10 cgc ttc cta ttc atg gag acc tac tat tac aag aac ccc tgc aag 624
Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Lys Asn Pro Cys Arg 200 205
cag gca ctg gag ctg gcc ccc gag gcc ctc ctt gcc ctg gcc acc ctc 672
Gln Ala Leu Glu Val Ala Pro Gly Ala Leu Leu Glu Gly Asn Leu 215 220
15 acc cag ctg tca ctc aag tac aac acc ctc act ctg ctg ccc cgc acc 720
Thr His Leu Ser Lys Tyr Asn Asn Leu Thr Val Val Pro Arg Asn 235 240
20 ctg cct tcc acc ctg gag tat ctg ctg ttc tac aac cgc atc gtc 768
Leu Pro Ser Ser 245 250
aaa ctg gag cct gag cag ctg gcc acc ctg acc ctc ctg cgt gtc ctc 816
Lys Leu Ala Pro Glu Asp Leu Ala Asn Leu Thr Ala Leu Arg Val Leu 265 270
25 gac ctg gag gaa aat tgc cgc cgc tgc gac cgc ctc acc ccc tgc 864
Asp Val Gly Gly Asn Cys Arg Arg Cys Asp His Ala Pro Asn Pro Cys 280 285
atg gag tgc cct gtc ctc acc ttc ccc cag cta cat ccc acc ttc acc 912
Met Glu Cys Pro Arg His Phe Pro Gln Leu His Pro Asp Thr Phe Ser 295 300
30 cag ctg gtc cgt ctt gaa gag ctg ctg gag cag gtc tct ctc tcc 960
His Leu Ser Arg Leu Glu Gly Leu Val Leu Lys Asp Ser Leu Ser 310 315
35 tgc ctg aat gcc agt tgc ttc gtc gag gaa aac ctc cga gtc ctg 1008
Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu 320 325
gac ctg agt gag aac ttc ctc tcc aag tgc acc act aac acc aag gcc 1056
Asp Leu Ser Glu Asn Phe Leu Tyr Lys Cys Ile Thr Lys Thr Lys Ala 330 335
40 ttc cag ggc cta aca cag ctg cgc aag ctt aac ctg tcc ttc aat tcc 1104
Phe Gln Gly Leu Thr Gln Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr 340 345
caa aag gag gtc tct ttt gcc ctc ctg tct ctg gcc cct tcc ttc gag 1152
Gln Lys Arg Val Ser Phe Ala Ala His Leu Ser Leu Ala Pro Ser Phe Gly 350 355
45 gac ctg gtc gcc ctg aag gag ctg gac cag gcc acc ttc ttc ctc cgc 1200
Ser Leu Val Ala Leu Lys Glu Leu Asp Met His Gly Ile Phe Phe Arg 360 365
tca ctc aat gag acc acc ctc cgc cga ctg gcc cgc ctg ccc atg ctc 1248
50 Ser Leu Asp Glu Thr Thr Leu Arg Pro Leu Ala Arg Leu Pro Met Leu 370 375
cag act ctg cgt ctg cag atg aac ttc acc aac cag gcc cgc ctc ggc 1296

Gln Thr Leu Arg Leu Gln Met Asn Phe Ile Asn Gln Ala Gln Leu Gly 420 430
atc ttc agc gcc ttc ctc ctg cgc tac ggc ctg ctg tgc gac aac 1344
11e Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn 435 440
5 cgc atc agc gga gct tgc gag ctg aca gcc acc atg gag gga gca gat 1392
Arg Ile Ser Gly Ala Ser Glu Leu Thr Ala Thr Met Gly Ala Asp 445 450
10 gag gag aag gtc tgc ctg cag cct gag gcc ctt gtc ctg gcc cca 1440
Gly Gly Glu Lys Val Trp Leu Gln Pro Gly Asp Leu Ala Pro Ala Pro 455 460
gag gac act ccc acc tct gaa gac ttc agc ccc aac tgc acc ctc ctc 480
Val Asp Thr Pro Ser Ser Glu Asp Phe Arg Pro Asn Cys Ser Thr Leu 485 490
15 aac ttc acc tgc gtc ctg tca cgg aac aac ctg gtc acc gtc cgc cgc 1536
Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Gln Pro 500 505
gag atg ttt gcc cag ctc tgc cag ctg gtc ctg gtc gtc gtc cgc 1584
Glu Met Phe Ala Gln Leu Ser His Leu Gln Cys Leu Arg Leu Ser His 510 515
20 aac tgc atc ctg cga gca gtc aat ggc tcc cag ttc ctg ctg cgc acc 1632
Asn Cys Ile Ser Gln Ala Val Asn Gly Ser Gln Phe Leu Pro Leu Thr 520 525
ggt ctg cag gtc cta gac ctg tcc ccc aat aag ctg gtc ctc tac ccc 1680
Gly Leu Gln Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His 530 535
25 gag cag ttc tgc agc gaa cta cca cga ctg gag gtc ctg gac ctc acc 1728
Glu His Ser Phe Thr Glu Leu Pro Arg Leu Glu Ala Leu Asp Leu Ser 540 545
30 ttc aac acc ggc ccc ttt ggc atg ctg gag ggc gtc gtc aac ttc acc 1776
Tyr Asn Ser Gln Pro Phe Gly Met Gln Gly Val Gly His Asn Phe Ser 550 555
ttc gtc gtc ctc cgc acc ctg cgc cgc ctc cgc ctg gtc cgc aac 1824
Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn 560 565
35 aac atc cag acc cca gtc tcc cag cag ctc tgc agt agt tgc ctg cgc 1872
Asn Ile His Ser Gln Val Ser Gln Gln Leu Cys Ser Thr Ser Leu Arg 570 575
gac ctg gac ttc acc ggc aat gca ctg ggc cct atg tgc gcc gag gga 1920
Ala Leu Asp Phe Ser Gly Asn Ala Leu Gly His Met Trp Ala Gly Gly 580 585
gac ctc tat ctg cgc ttc cca ggc ctg acc ggt tgc atc tgc ctg 1968
Asp Leu Tyr Leu His Phe Phe Gln Gly Leu Ser Gly Leu Ile Trp Leu 590 595
40 gag tgc tcc cag acc cgc ctg ccc acc ctc ctg ccc aac acc ctg cgc 2016
Asp Leu Ser Gln Asn Arg Leu His Thr Leu Leu Pro Gln Thr Leu Arg 600 605
aac ctc ccc aag acc cta cag gtc ctg cgt ctc cgt gac aat tac ctg 2064
Asn Leu Pro Lys Ser Leu Gln Val Leu Arg Leu Asp Asn Tyr Leu 610 615
50 ggc ttc ttt aag tgc tgc acc ctc ctc ctg ccc aac ctg gaa gtc 2112
Ala Phe Phe Lys Trp Ser Leu His Phe Leu Pro Lys Leu Glu Val 620 625

30

690 ctc cag gca gga aac cag cty aag gcc cty acc aac ggc acc cty 2160
Leu Asp Leu Ala Gly Asn Glu Leu Lys Ala Thr Asn Gly Ser Leu
705 710 715 720
5 gct gct ggc acc cag ctc cag gga cty gat gtc acc ggc acc gtc 2208
Pro Ala Gly Thr Arg Leu Arg Leu Asp Val Ser Cys Asn Ser Ile
725 730 735
gpc ttc gty gcc ccc ggc ttc ttc ttc aag gcc aag gaa cty gaa gaa 2256
Ser Phe Val Pro Gly Phe Ser Lys Ala Lys Glu Leu Arg Glu
740 745 750
ctc aac ctt acc ggc aac gcc ctc aag aca gty gac cac ttc ttc ttc 2304
Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe
755 760 765
ggg ccc cty gga ggt gcc cty cca ata cta gat gta acc ggc acc cct 2352
15 Gly Pro Leu Ala Ser Ala Leu Glu Ile Leu Asp Val Ser Ala Asn Pro
770 775 780
ctc cac tgc gcc tct ggg gca gcc ttt atg gac ttc cty cty gaa gty 2400
Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Glu Val
785 790 795
cag gct ggc gty ccc ggt cty ccc acc cag gga aag tct ggc acc cag 2448
Gln Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro
805 810 815
ggc cag ctc cag ggc ctc acc ttt gca cag gac cty gcc ctc tgc 2496
25 Gly Glu Leu Glu Gly Leu Ser Ile Phe Ala Glu Asp Leu Leu Cys
820 825 830
ctg gat gaa gcc ctc ttc tgc gaa tct ttc gtc ctc tgc cty gct gct 2544
Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu Ala
835 840 845
gty gct cty gcc cty ggt gty cct atg cty cat cac ctc tct ggc tgg 2592
30 Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly Trp
850 855 860
gpc ctc tgc tac tgc ttc cac cty gcc cty gcc tgg ctt ccc tgc cga 2640
Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg
865 870 875
ggg cga caa agt ggg cga gat gaa gct gcc cty ccc tac gat gcc ttc 2688
35 Gly Arg Glu Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala Phe
885 890 895
gty gtc ttc gac aaa aag cag gcc gca gty gca gac tgg gty tac aac 2736
40 Val Val Phe Asp Lys Thr Glu Ser Ala Val Ala Asp Trp Val Tyr Asn
900 905 910
gag ctt cga gga cag cty gaa gaa tgc cgt ggg gcc tgc gca ctc cgc 2784
Glu Leu Arg Gly Glu Leu Glu Glu Cys Arg Gly Arg Trp Ala Leu Arg
915 920 925
ctg tgc cty gaa gaa gcc pac tgg cty cct ggc aaa acc ctc ttc gaa 2832
45 Leu Cys Leu Glu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Glu
930 935 940
aac cty tgg gcc tgc gtc tat gcc gcc acc aag acc cty ttc gty cty 2880
Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu
945 950 955
gac cac aag gcc cag gtc aat gct ctc tgc gcc gcc acc ttc cty cty 2928
60 Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu
965 970 975

31

ggc cag cag gcc cty gaa gac gcc aag gcc gtc gty gty cty gty 2976
Ala Glu Glu Arg Leu Leu Glu Asp Arg Lys Asp Val Val Leu Val
980 985 990
atc cty gcc cct gcc gcc gcc gcc ctc ctc gtc tat gty cgg cty gcc cag 3024
5 Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Val Arg Leu Arg Glu
995 1000 1005
gpc ctc tgc gcc cag agt gtc ctc ctc tgc ccc cac cag ccc agt ggt 3072
Arg Leu Cys Arg Glu Ser Val Leu Leu Trp Pro His Glu Pro Ser Gly
1010 1015 1020
cag gcc gcc ttc tgg gcc cty gcc atg gcc cty acc agg gcc aac 3120
10 Gln Arg Ser Phe Trp Ala Glu Leu Gly Met Ala Leu Thr Arg Asp Asn
1025 1030 1035
cac cac ttc tat aac cgg acc ttc tgc cag gga ccc acc gcc gaa 3165
15 His His Phe Tyr Asn Arg Asn Phe Cys Glu Gly Pro Thr Ala Glu
1045 1050 1055
<210> 2
<211> 1055
<212> PRT
20 <213> Homo sapiens
<400> 2
Met Pro Met Lys Trp Ser Gly Trp Arg Trp Ser Trp Gly Pro Ala Thr
1 5 10 15
25 His Thr Ala Leu Pro Pro Pro Glu Gly Phe Cys Arg Ser Ala Leu His
20 25 30
Pro Leu Ser Leu Leu Val Glu Ala Ile Met Leu Ala Met Thr Leu Ala
35 40 45
Leu Gly Thr Leu Pro Ala Phe Leu Pro Cys Glu Leu Glu Pro His Gly
50 55 60
Leu Val Asn Cys Asn Trp Leu Phe Leu Lys Ser Val Pro His Phe Ser
65 70 75
Met Ala Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn
80 85 90
35 Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu
100 105 110
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Pro Val Gly Leu Ser Pro
115 120 125
Met His Phe Pro Cys His Met Thr Ile Glu Pro Ser Thr Phe Leu Ala
130 135 140
40 Val Pro Thr Leu Glu Glu Leu Asn Leu Ser Tyr Asn Asn Ile Met Thr
145 150 155
Val Pro Ala Leu Pro Lys Ser Leu Ile Ser Leu Ser Leu Ser His Thr
165 170 175
45 Asn Ile Leu Met Leu Asp Ser Ala Ser Leu Ala Gly Leu His Ala Leu
180 185 190
Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Tyr Lys Asn Pro Cys Arg
195 200 205
Gln Ala Leu Glu Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu
210 215 220
50 Thr His Leu Ser Leu Lys Tyr Asn Asn Leu Thr Val Val Pro Arg Asn
225 230 235 240

32

Leu Pro Ser Ser Leu Glu Tyr Leu Leu Leu Ser Tyr Asn Arg Ile Val 250
 Lys Leu Ala Pro Glu Asp Leu Ala Asn Leu Thr Ala Leu Arg Val Leu 255
 5 Asp Val Gly Gly Asn Cys Arg Arg Cys Asp His Ala Pro Asn Pro Cys 260
 Met Glu Cys Pro Arg His Phe Pro Glu Leu His Pro Asp Thr Phe Ser 265
 His Leu Ser Arg Leu Glu Gly Leu Val Leu Lys Asp Ser Ser Leu Ser 270
 10 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415
 Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu
 Asp Leu Ser Glu Asn Phe Leu Tyr Lys Cys Ile Thr Lys Thr Lys Ala
 Phe Glu Gly Leu Thr Glu Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr
 Glu Lys Arg Val Ser Phe Ala His Leu Ser Leu Ala Pro Ser Phe Gly
 Ser Leu Val Ala Leu Lys Glu Leu Asp Met His Gly Ile Phe Phe Arg
 20 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655
 Ser Leu Asp Glu Thr Thr Leu Arg Pro Leu Ala Arg Leu Pro Met Leu
 Glu Thr Leu Arg Leu Glu Met Asn Phe Ile Asn Glu Ala Glu Leu Gly
 Arg Ile Ser Gly Ala Ser Glu Leu Thr Ala Thr Met Gly Glu Ala Asp
 Gly Gly Glu Lys Val Trp Leu Glu Pro Gly Asp Leu Ala Pro Ala Pro
 Val Asp Thr Pro Ser Ser Glu Asp Phe Arg Pro Asn Cys Ser Thr Leu
 Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Glu Pro
 35 Glu Met Phe Ala Glu Leu Ser His Leu Glu Cys Leu Arg Leu Ser His
 Asn Cys Ile Ser Glu Ala Val Asn Gly Ser Glu Phe Leu Pro Leu Thr
 Gly Leu Glu Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His
 40 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655
 Glu His Ser Phe Thr Glu Leu Pro Arg Leu Glu Ala Leu Asp Leu Ser
 Tyr Asn Ser Glu Pro Phe Gly Met Glu Val Gly His Asn Phe Ser
 Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn
 Asn Ile His Ser Glu Val Ser Glu Glu Leu Cys Ser Thr Ser Leu Arg
 Ala Leu Asp Phe Ser Gly Asn Ala Leu Gly His Met Trp Ala Glu Gly
 50 625 630 635 640 645 650 655
 Asp Leu Tyr Leu His Phe Glu Glu Ser Gly Leu Ile Trp Leu

SHEDDING-47C.....0155386A.1..

33

Asp Leu Ser Glu Asn Arg Leu His Thr Leu Leu Pro Glu Thr Leu Arg 660
 Asn Leu Pro Lys Ser Leu Glu Val Val Arg Leu Arg Asp Asn Tyr Leu 665
 5 Ala Phe Phe Lys Trp Ser Leu His Phe Leu Pro Lys Leu Glu Val 670
 Leu Asp Leu Ala Gly Asn Glu Leu Lys Ala Leu Thr Asn Gly Ser Leu 675
 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995 1000 1005 1010 1015 1020 1025 1030 1035 1040 1045 1050 1055
 Pro Ala Gly Thr Arg Leu Arg Leu Asp Val Ser Cys Asn Ser Ile
 Ser Phe Val Ala Pro Gly Phe Phe Ser Lys Ala Lys Glu Leu Arg Glu
 Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe
 15 Gly Pro Leu Ala Ser Ala Leu Glu Ile Leu Asp Val Ser Ala Asn Pro
 Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Glu Val
 Glu Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro
 20 Gly Glu Leu Glu Gly Leu Ser Ile Phe Ala Glu Asp Leu Arg Leu Cys
 Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu Ala
 Val Val Leu Gly Leu Gly Val Pro Met Leu His Leu Cys Gly Trp
 Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg
 Gly Arg Glu Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala Phe
 30 Val Val Phe Asp Lys Thr Glu Ser Ala Val Ala Asp Trp Val Tyr Asn
 Glu Leu Arg Gly Glu Leu Glu Cys Arg Gly Arg Trp Ala Leu Arg
 35 Leu Cys Leu Glu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Glu
 Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu
 Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu
 40 Ala Glu Glu Arg Leu Leu Glu Asp Arg Lys Asp Val Val Leu Val
 Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Val Arg Leu Arg Glu
 Arg Leu Cys Arg Glu Ser Val Leu Trp Pro His Glu Pro Ser Gly
 Glu Arg Ser Phe Trp Ala Glu Leu Gly Met Ala Leu Thr Arg Asp Asn
 His His Phe Tyr Asn Arg Asn Phe Cys Glu Gly Pro Thr Ala Glu
 50 1050 1055

SHEDDING-47C.....0155386A.1..

34

Seq. Id. no. 3

atg ggt ttc tgc cgc agc gcc ctg cac
met gty phe cys arg ser ala leu his
5
5 ccc ctg tct ctc ctg ctg cag gcc atc atg ctg gcc atg acc ctg gcc 75
Pro Leu Ser Leu Leu Val Gln Ala Ile Met Leu Ala Met Thr Leu Ala
10
ctg ggc acc ttg cct gcc ttc cta ccc tct gag ctc cag ccc cag gcc 123
Leu Gly Thr Leu Pro Ala Phe Leu Pro Cys Gly Leu Gln Pro His Gly
30
ctg ctg aac tgc aac tgg ctg ttc ctg aag tct gty ccc cag ttc ttc 40
Leu Val Asn Cys Asn Trp Leu Phe Leu Lys Ser Val Pro His Phe Ser
45
atg gca gca ccc cgt gcc aat gtc acc gcc cct tcc ttg tcc tcc aac 55
Met Ala Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn
60
cgc atc cag cac ctc cat gat tct gac ttt gcc cag ctg ccc acc ctg 70
Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu
75
cgc cat ctc aac ctg aag tgg aac tgc cgc ggt gcc ctg ctg acc ccc 85
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Pro Val Gly Leu Ser Pro
90
atg cag ttc ccc tgc cag atc atc gag ccc gcc acc ttc ttg gct 105
Met His Phe Pro Cys His Met Thr Ile Gln Pro Ser Thr Phe Leu Ala
110
gty ccc acc ctg gaa gag cta aac ctg agc tgc aac aac atc atg act 120
Val Pro Thr Leu Gln Gln Leu Asn Leu Ser Tyr Asn Asn Ile Met Thr
125
gty cct ggc ctg ccc aaa tcc ctc ata tcc ctg tcc ctc acc cct acc 135
Val Pro Ala Leu Pro Lys Ser Leu Ile Ser Leu Ser Leu Ser His Thr
140
aac atc ctg atg cta gcc tct gcc gcc ctc gcc gcc ctg cat gcc ctg 150
Asn Ile Leu Met Leu Asp Ser Ala Ser Leu Ala Gly Leu His Ala Leu
155
cgc ttc cta ttc atg gac gcc aac tct tat tac aag aac ccc tgc agc 165
Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Tyr Lys Asn Pro Cys Arg
170
cag gca ctg gag ctg gcc ccc ggc tcc ctc ctc gcc ctg gcc aac ctc 185
Gln Ala Leu Gln Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu
190
acc cag ctg tca ctc aag tac aac acc ctc act gty gty ccc gcc aac 200
Thr His Leu Ser Leu Lys Tyr Asn Asn Leu Thr Val Pro Arg Asn
205
ctg ctc acc ctg gag tat ctg ctg ttg ttc tac aac gcc atc gtc 215
Leu Pro Ser Ser Leu Gln Tyr Trp Leu Leu Leu Ser Tyr Asn Arg Ile Val
220
aaa ctg ggc cct gag gcc ctg gcc aat ctg acc gcc ctg ggc ctg ctc 230
Lys Leu Ala Pro Gln Asp Leu Ala Asn Leu Thr Ala Leu Arg Val Leu
235
gat gty ggc gga aat tgc gcc gcc tgc gcc cag ccc acc ccc tgc 245
Asp Val Gly Gly Asn Cys Arg Cys Asp His Ala Pro Asn Pro Cys
250
255

35

atg gag tgc cct cgt cac ttc ccc cag cta cat ccc gtt acc ttc agc 843
Met Gln Cys Pro Arg His Phe Pro Gln Leu His Pro Asp Thr Phe Ser
270
cac ctg agc cgt ctt gaa gcc ctg gty ttg aag gac agt tct ctc tcc 891
His Leu Ser Arg Leu Gln Gly Leu Val Leu Lys Asp Ser Ser Leu Ser
285
tgg ctg aat gcc agt tgg ttc cgt ggg ctg gga aac ctc cag gty ctg 939
Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu
300
gac ctg agt gag aac ttc ctc tac aaa tgc atc act aaa acc aag gcc 987
Asp Leu Ser Gln Asn Phe Leu Tyr Lys Cys Ile Thr Lys Thr Lys Ala
315
ttc cag gcc cta aca cag ctg cgc aag ctt aac ctg tcc ttc aat tac 1035
Phe Gln Gly Leu Thr Gln Leu Arg Lys Leu Asn Leu Ser Asn Tyr
330
caa aag aag gty ttc ttc gcc cag ctg tct ctg gcc cct tcc ttc ggg 1083
Gln Lys Arg Val Ser Phe Ala His Leu Ser Leu Ala Pro Ser Phe Gly
350
agc ctg gtc gcc ctg aag gag ctg gcc atg cag gcc atc ttc ttc cgc 1131
Ser Leu Val Ala Leu Lys Gln Leu Asp Met His Gly Ile Phe Phe Arg
365
tca ctc gat gag acc aag ctc cgg cca ctg gcc cgc ctg acc atg ctg 1179
Ser Leu Asp Gln Thr Thr Leu Arg Pro Leu Ala Arg Leu Pro Met Leu
380
cag act ctg ggt ctg cag atg aac ttc atc aac cag gcc cag ctc gcc 1227
Gln Thr Leu Arg Leu Gln Met Asn Phe Ile Asn Gln Ala Gln Leu Gly
395
atc ttc aag gcc ttc cct gcc ctg cgc tac gty gcc ctg tgc gcc aac 1275
Ile Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn
410
cgc atc agc gga gtt tgc gag ctg aca gcc acc atg ggg gag cta gtt 1323
Arg Ile Ser Gly Ala Ser Gln Leu Thr Ala Thr Met Gly Gln Ala Asp
430
gga ggg gag aag gtc tgg ctg cag cct ggg gcc ctt gct cgc gcc cca 1371
Gly Gly Gln Lys Val Trp Leu Gln Pro Gly Asp Leu Ala Pro Ala Pro
445
gty gac act ccc agc tct gaa gac ttc agc ccc aac tgc agc acc ctc 1419
Val Asp Thr Pro Ser Ser Gln Asp Phe Arg Pro Asn Cys Ser Thr Leu
460
aac ttc acc ttg gat ctg tca cgg aac aac ctg gty acc gty cag cga 1467
Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Gln Pro
475
gag atg ttc gcc cag ctc ttg cag ctg cag tgc ctg ctg agc cag 1515
Glu Met Phe Ala Gln Leu Ser His Leu Gln Cys Leu Arg Leu Ser His
490
aac tgc atc tgc cag gtc atc ggc tcc cag ttc ctg cgc ctg acc 1563
Asn Cys Ile Ser Gln Ala Val Asn Gly Ser Gln Phe Leu Pro Leu Thr
510
ggt ctg cag gta cta gac ctg tcc cag aat aag ctg gcc ctc ttc cag 1611
Gly Leu Gln Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His
525
gag cac tca ttc atg gag cta cca cga ctg gag gcc ctg gcc ctc gcc 1669

38

Met Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn 45 55
60 65 70
Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu
75 80 85
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Val Gly Leu Ser Pro
90 95 100
Met His Phe Pro Cys His Met Thr Ile Glu Pro Ser Thr Phe Leu Ala
110 115 120
Val Pro Thr Leu Glu Glu Leu Asn Leu Ser Tyr Asn Asn Ile Met Thr
125 130 135
Val Pro Ala Leu Pro Lys Ser Leu Ile Ser Leu Ser Leu Ser His Thr
140 145 150
Asn Ile Leu Met Leu Asp Ser Ala Ser Leu Ala Gly Leu His Ala Leu
155 160 165
Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Tyr Lys Asn Pro Cys Arg
170 175 180
Gln Ala Leu Glu Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu
190 195 200
Thr His Leu Ser Leu Lys Tyr Asn Asn Leu Thr Val Val Pro Arg Asn
202 210 215
Leu Pro Ser Ser Leu Glu Tyr Leu Leu Ser Tyr Asn Arg Ile Val
220 225 230
Lys Leu Ala Pro Glu Asp Leu Ala Asn Leu Thr Ala Leu Arg Val Leu
235 240 245
Asp Val Gly Gly Asn Cys Arg Cys Asp His Ala Pro Asn Pro Cys
250 255 260
Met Glu Cys Pro Arg His Phe Pro Gln Leu His Pro Asp Thr Phe Ser
270 275
His Leu Ser Arg Leu Glu Gly Leu Val Lys Asp Ser Leu Ser
285 290 295
Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu
300 305 310
Asp Leu Ser Glu Asn Phe Leu Tyr Lys Cys Ile Thr Lys Thr Lys Ala
315 320 325
Phe Gln Gly Leu Thr Gln Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr
330 335 340 345
Gln Lys Arg Val Ser Phe Ala His Leu Ser Leu Ala Pro Ser Phe Gly
350 355 360
Ser Leu Val Ala Leu Lys Glu Leu Asp Met His Gly Ile Phe Phe Arg
365 370 375
Ser Leu Asp Glu Thr Thr Leu Arg Pro Leu Ala Arg Leu Pro Met Leu
380 385 390
Gln Thr Leu Arg Leu Gln Met Asn Phe Ile Asn Gln Ala Gln Leu Gly
395 400 405
Ile Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn
410 415 420 425
Arg Ile Ser Gly Ala Ser Glu Leu Thr Ala Thr Met Gly Glu Ala Asp
430 435 440
Gly Gly Glu Lys Val Trp Leu Gln Pro Gly Asp Leu Ala Pro Ala Pro
445 450 455
Val Asp Thr Pro Ser Ser Glu Asp Phe Arg Pro Asn Cys Ser Thr Leu

BIOCHECK-ARG-GENEWAL.LJ>

39

Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Gln Pro 450 455 470
475 480 485
Glu Met Phe Ala Gln Leu Ser His Leu Gln Cys Leu Arg Leu Ser His
490 495 500
Asn Cys Ile Ser Gln Ala Val Asn Gly Ser Gln Phe Leu Pro Leu Thr
510 515 520
Gly Leu Gln Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His
525 530 535
Glu His Ser Phe Thr Glu Leu Pro Arg Leu Glu Ala Leu Asp Leu Ser
540 545 550
Tyr Asn Ser Gln Pro Phe Gly Met Gln Gly Val Gly His Asn Phe Ser
555 560 565
Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn
570 575 580
Asn Ile His Ser Gln Val Ser Gln Gln Leu Cys Ser Thr Ser Leu Arg
590 595
Ala Leu Asp Phe Ser Gly Asn Ala Leu Gly His Met Trp Ala Glu Gly
605 610 615
Asp Leu Tyr Leu His Phe Phe Gln Gly Leu Ser Gly Leu Ile Trp Leu
620 625 630
Asp Leu Ser Gln Asn Arg Leu His Thr Leu Leu Pro Gln Thr Leu Arg
635 640 645
Asn Leu Pro Lys Ser Leu Gln Val Leu Arg Leu Arg Asp Asn Tyr Leu
650 655 660
Ala Phe Phe Lys Trp Trp Ser Leu His Phe Leu Pro Lys Leu Glu Val
670 675 680
Leu Asp Leu Ala Gly Asn Gln Leu Lys Ala Leu Thr Asn Gly Ser Leu
685 690 695
Pro Ala Gly Thr Arg Leu Arg Arg Leu Asp Val Ser Cys Asn Ser Ile
700 705 710
Ser Phe Val Ala Pro Gly Phe Ser Lys Ala Lys Glu Leu Arg Glu
715 720 725
Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe
730 735 740 745
Gly Pro Leu Ala Ser Ala Leu Gln Ile Leu Asp Val Ser Ala Asn Pro
750 755 760
Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Val
765 770 775
Gln Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro
780 785 790
Gly Gln Leu Gln Gly Leu Ser Ile Phe Ala Gln Asp Leu Arg Leu Cys
795 800 805
Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu Ala
810 815 820 825
Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly Tyr
830 835 840
Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg
845 850 855
Gly Arg Gln Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala Phe
860 865 870
Val Val Phe Asp Lys Thr Gln Ser Ala Val Ala Asp Trp Val Tyr Asn

BIOCHECK-ARG-GENEWAL.LJ>

CLAIMS

1. An isolated Toll-like-receptor polypeptide consisting essentially of

- (i) the amino acid sequence SEQ ID NO: 2;
 (ii) a variant thereof which has immunomodulatory activity; or
 (iii) a fragment of (i) or (ii) which has immunomodulatory activity.

2. A polypeptide according to claim 1 wherein the variant (ii) has at least 70% identity to the amino acid sequence of SEQ ID NO: 2.

3. A polypeptide according to claim 2 which has at least 70% identity to the amino acid sequence of SEQ ID NO: 2 over a region of at least 70% of the full-length sequence provided by SEQ ID No.1 and exhibits toll-like receptor functionality.

4. A polypeptide according to claim 1 or claim 2 wherein the variant (ii) has at least 95% identity to the amino acid sequence of SEQ ID NO: 2.

5. A polypeptide according to claim 2 which has at least 95% identity to the amino acid sequence of SEQ ID NO: 2 over a region of at least 60 contiguous amino acids and exhibits toll-like receptor functionality.

6. A polypeptide according to claim 1 wherein the fragment (iii) is a peptide of up to 150 amino acids in length and exhibits toll-like receptor functionality.

7. A polynucleotide encoding a polypeptide according to any one of claims 1-3.

8. A polynucleotide encoding a Toll-like receptor polypeptide which has immunomodulatory activity, which polynucleotide consists essentially of:

- (a) the nucleic acid sequence of SEQ ID NO: 1;
 (b) a sequence complementary thereto;

875 Glu Leu Arg Gly Gln Leu Glu Cys Arg Gly Arg Trp Ala Leu Arg 885
 890 895 900 905
 Leu Cys Leu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Glu
 910 915 920
 Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu
 925 930 935
 Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu
 940 945 950
 Ala Gln Gln Arg Leu Leu Glu Asp Arg Lys Asp Val Val Leu Val
 955 960 965
 Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Val Arg Leu Arg Gln
 970 975 980 985
 Arg Leu Cys Arg Gln Ser Val Leu Leu Trp Pro His Gln Pro Ser Gly
 990 995 1000
 Gln Arg Ser Phe Trp Ala Gln Leu Gly Met Ala Leu Thr Arg Asp Asn
 1005 1010 1015
 His His Phe Tyr Asn Arg Asn Phe Cys Gln Gly Pro Thr Ala Glu
 1020 1025 1030

42

- (c) a sequence which hybridises under stringent conditions to a sequence as defined in (a) or (b);
 (d) a sequence that is degenerate as a result of the genetic code to a sequence as defined in (a), (b) or (c); or
 (e) a sequence having at least 60% identity to a sequence as defined in (a), (b), (c) or (d).

9. A polynucleotide according to claim 7 or claim 8 which is mRNA.

10. A polynucleotide according to claim 7 or claim 8 which is DNA.

11. A polynucleotide according to claim 7 or claim 8 which is cDNA.

12. An isolated toll-like receptor polypeptide which is obtainable by expression *in vivo* or *in vitro* of a polynucleotide according to claim 7 or claim 8.

13. A polypeptide according to claim 12 which has the structural features conserved amongst toll-like receptors.

14. An expression vector comprising a polynucleotide sequence according to any one of claims 7 to 11, which is capable of expressing a polypeptide according to any one of claims 1 to 3 or claim 12.

15. An expression vector according to claim 14 which is a plasmid, phage or viral vector.

16. A host cell comprising an expression vector according to claim 14 or claim 15.

43

17. A polyclonal or monoclonal antibody, or a chimera or fragment thereof, which is specific for a polypeptide according to any one of claims 1 to 3.

18. A method for identification of a compound that modulates Toll-like receptor activity, which method comprises contacting a polypeptide according to any one of claims 1 to 3 or claim 12 with a test substance and monitoring for immunomodulatory activity.

19. A compound which modulates Toll-like receptor activity and which is identifiable by a method according to claim 18.

20. A compound according to claim 19 which is a peptide or polypeptide.

21. A compound according to claim 19 which is an oligonucleotide or polynucleotide.

22. A method of treating a subject having an inflammatory or cardiovascular disorder, systemic infection or autoimmune disease that is responsive to Toll-like receptor modulation, which method comprises administering to said subject an effective amount of a compound according to any one of claims 19 to 21 or an antibody according to claim 17.

23. A method according to claim 22 wherein the disorder is a viral, fungal or bacterial infection, asthma, rhinitis, chronic obstructive pulmonary disease (COPD), emphysema, an inflammatory bowel disease such as ulcerative colitis or Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, Alzheimer's disease, atherosclerosis, Multiple Sclerosis, diabetes or septic shock syndrome associated with systemic infection involving gram positive or gram negative bacteria.

44

24. A polypeptide comprising a fragment or variant of SEQ ID NO: 2, which is capable of inhibiting the activity of TLR9 having the amino acid sequence of SEQ ID NO: 2 or a functional variant thereof, for use in the treatment of an immune or inflammatory disorder.

5

25. A polypeptide according to any one of claims 1 to 3 or claim 12, a polynucleotide according to claim 7 or claim 8 or a compound according to claim 19 for use as an adjuvant.

10 26. The use of a compound according to claim 19 in the manufacture of a medicament for the treatment of an immune or inflammatory disorder.

15 27. The use of a polypeptide comprising a fragment or variant of SEQ ID NO: 2, which is capable of inhibiting the activity of TLR9 having the amino acid sequence of SEQ ID NO: 2 or a functional variant thereof, in the manufacture of a medicament for the treatment of an immune or inflammatory disorder.

20 28. The use of a polypeptide according to any one of claims 1 to 3 or claim 12, a polynucleotide according to claim 7 or claim 8 or a compound according to claim 19 in the manufacture of an adjuvant or vaccine formulation.

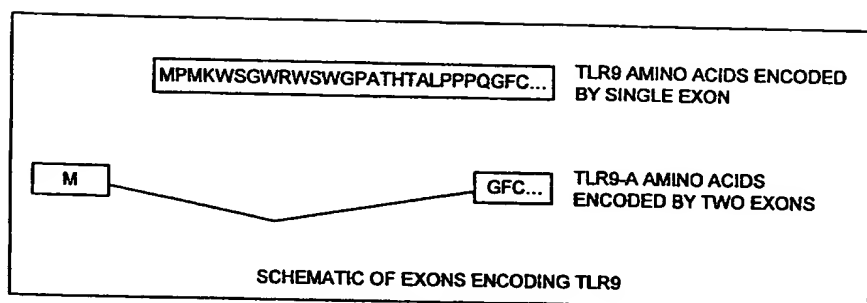
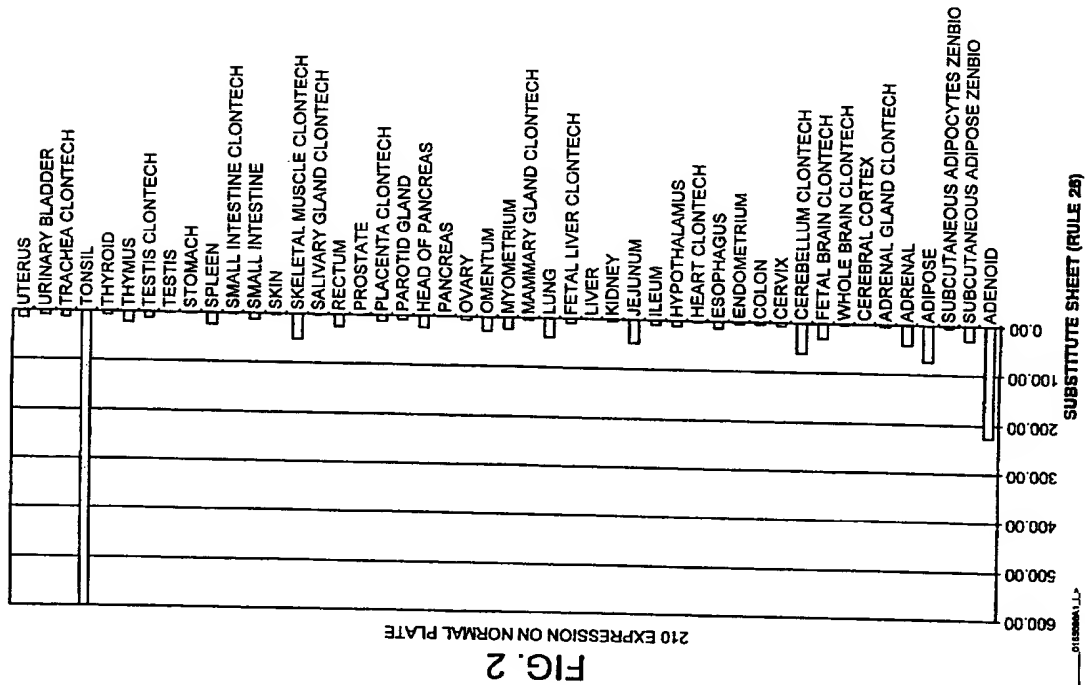
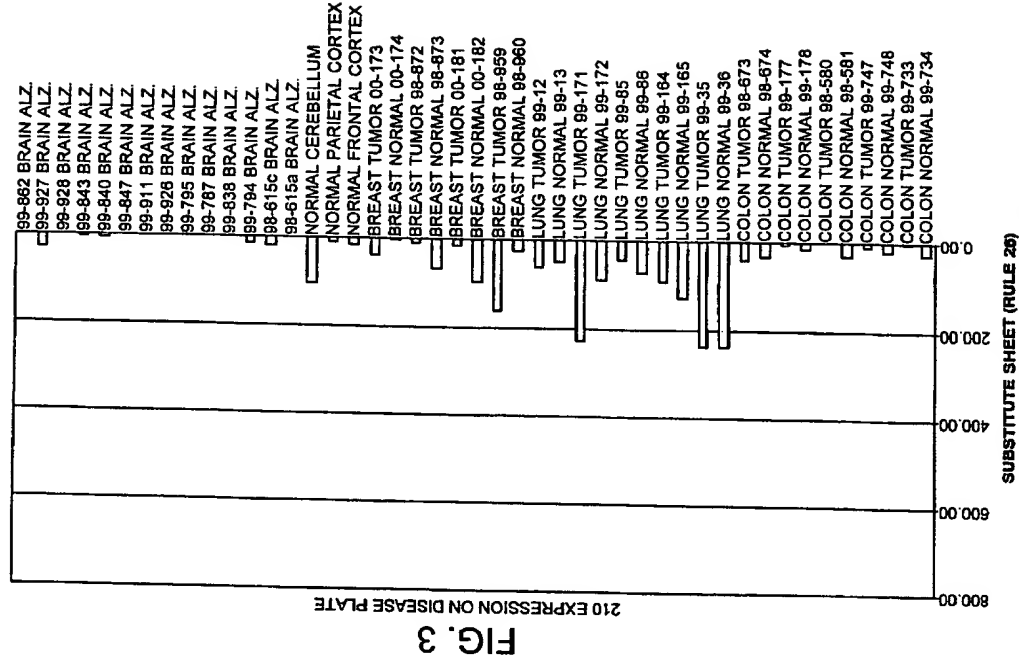


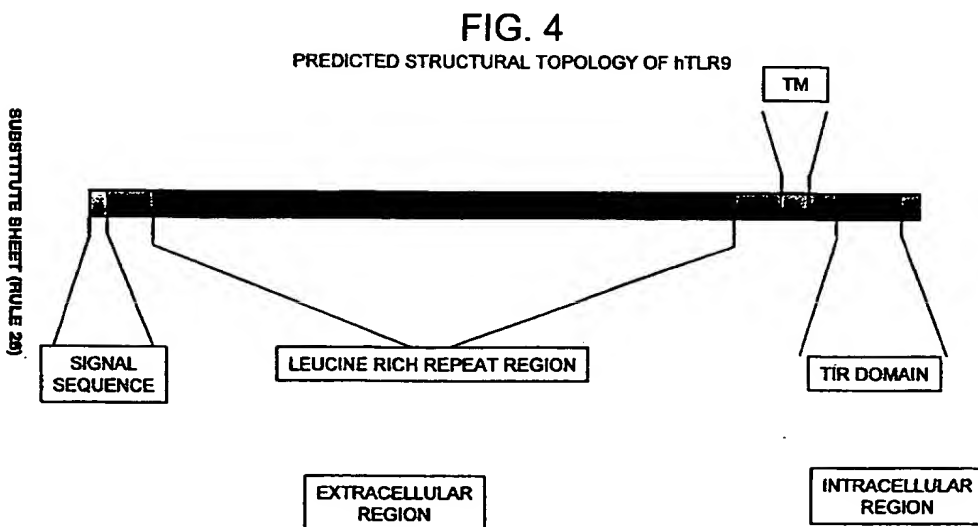
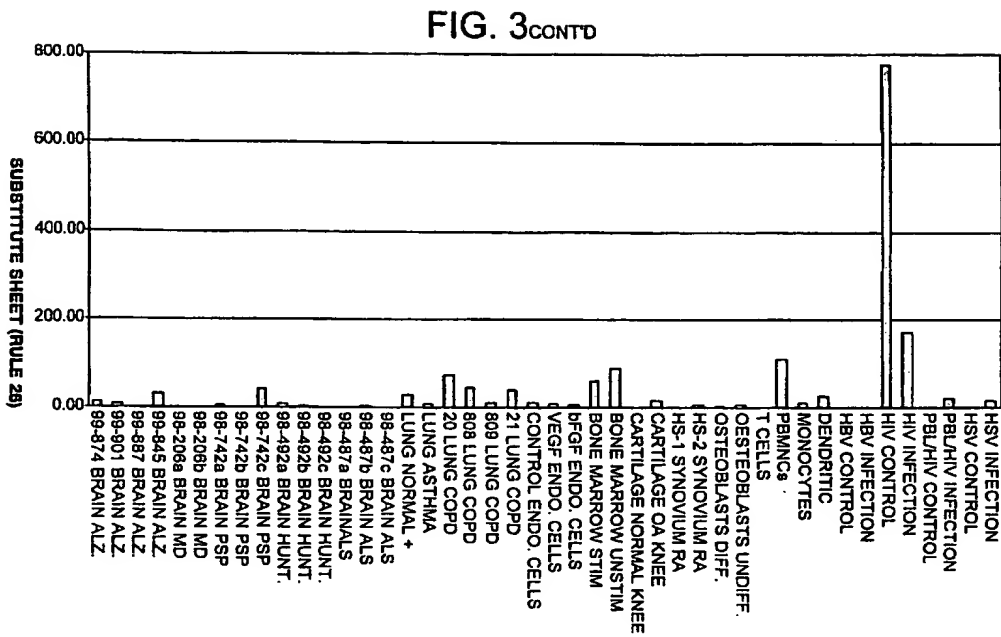
FIG. 1

2 / 5



3 / 5





INTERNATIONAL SEARCH REPORT		Intern. Appl. No. PCT/GB 01/00299	
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C12N15/63 C12N15/10 C07K16/28 A61K31/7088 A61K38/17 A61K48/00 C07K14/705		G01N33/68 A61K48/00	
According to International Patent Classification (IPC) or to both related classification and IPC B. FIELDS SEARCHED IPC 7 C12N 601N C07K A61K			
Documentation searched other than minimum documentation is the extent that such documents are indicated in this field searched			
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EMBL, STRAND, EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE, CHEM ABS Data			
C. DOCUMENTS CONSIDERED TO BE RELEVANT Category * Citation of document, with indication, where appropriate, of the relevant passage			
X	WO 98 50547 A (SCHERING CORP) 12 November 1998 (1998-11-12) 99.8% identity in 1008 bp overlap between SEQ ID NO 33 of W09850547 and SEQ ID NO 1 99.7% identity in 336 amino acids overlap between SEQ ID NO 34 of W09850547 and SEQ ID NO 2 page 56, line 23 -page 60, line 14; claims 1-16, examples 1-11		1-18, 22-25, 27,28
			Patent to claim No.
X	Further documents are listed in the continuation of box C.	X	Patent family members are listed in annex.
* Special categories of cited documents : "X" document affecting the patent side of the art which is not considered to be of particular relevance "Y" document published prior to the international filing date "Z" earlier document but published on or after the international filing date "1" document which is prior art for the purposes of the international search but which is not prior art for the purposes of the national or other special (main) (as specified) "2" document relating to an oral disclosure, use, exhibition or other information "3" document published prior to the international filing date but later than the priority date claimed			
Date of the actual completion of the international search			
28 June 2001		11/07/2001	
Name and mailing address of the ISA European Patent Office, P.O. Box 5518 Patenkaser 2 NL - 2005 HV Noordwijk, The Netherlands Fax: (31-73) 340-3015		Authorized officer Devlvyer, K	

INTERNATIONAL SEARCH REPORT		Serial Application No PCT/GB 01/00299
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		Relevant to claim No.
Category *	Citation of document, with indication, where appropriate, of the relevant passages	
X	<p>DATABASE EMBL 'Online! accession: AC006252, 29 December 1998 (1998-12-29) MUZYNI D ET AL: "Homo sapiens 3p21.1 contig 9 PAC RPC15-1157M23 (Roswell Park Cancer Institute Human PAC Library) complete sequence." XP002170407 100% identity in 3165 bp overlap (between positions 26803-29967) with SEQ ID NO 1 100% identity in 1055 amino acids overlap (between positions 26803-29967) with SEQ ID NO 2 (tfasta)</p>	7-11
P, X	<p>HEMI HIROAKI ET AL: "A Toll-like receptor recognizes bacterial DNA" NATURE/MACMILLAN JOURNALS LTD, LONDON, GB, vol 408, no. 6813, 7 December 2000 (2000-12-07), pages 740-746; XP002168474 ISSN: 0028-0836 cited in the application the whole document -& DATABASE EMBL 'Online! accession: AB045180, 13 December 2000 (2000-12-13) AKIRA S ET AL: "Homo sapiens TLR9 mRNA for toll-like receptor 9, complete cds." XP002170408 99.7% identity in 3113 bp overlap with SEQ ID NO 1 100% identity in 1031 amino acids overlap with SEQ ID NO 2</p>	1-18, 22-25, 27, 28

INTERNATIONAL SEARCH REPORT

Information on patent family members

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9850547	A	12-11-1998	AU 7175498 A BR 9808747 A CN 1263555 T EP 0980429 A NO 995458 A PL 336635 A SK 146589 A HU 0001462 A	27-11-1998 11-07-2000 16-08-2000 23-02-2000 08-11-1999 03-07-2000 11-07-2000 28-07-2000

Intern. Appl. No. PCT/88 01/00299

International Application No. PCT/88 01/00299

FURTHER INFORMATION CONTINUED FROM PCT/88/ 210

Continuation of Box I.2

Claims Nos.: 19-21,26; in part: 22,23,25,28

Claim 19 refers to a compound which modulates Toll-like receptor activity without giving a true technical characterization. Moreover, no such compounds are defined in the application. In consequence, the scope of said claim is ambiguous and vague, and its subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT).

No meaningful search can be carried out for such a purely speculative claim whose wording is, in fact, a mere recitation of the result to be achieved. The above comment also applies for claims 20,21 and 26; and in part for claims 22,23,25 and 28.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

Form PCT/ISA/210 (patent family annex) (July 1997)

ENCLOSURE -4PO 01182884 1,1

ENCLOSURE -4PO 01182884 1,1

THIS PAGE BLANK (USPTO)

Toll-like Receptor

Field of the invention

The present invention relates to a novel Toll-like receptor or a variant thereof. A variant may demonstrate Toll-like receptor activity such as activation of NF- κ B, or may inhibit Toll-like receptor activity.

Background of the invention

A family of human Toll-like receptors has been described in the literature. These receptors are termed Toll-like receptors in view of common structural features shared with a Drosophila Toll (dToll) receptor molecule which is involved in embryonic development. Toll and Toll-Like receptors are type I transmembrane proteins, with extracellular leucine rich repeat motifs and an intracellular signalling domain homologous to that of members of the interleukin 1 receptor superfamily.

Drosophila Toll also plays an important role in the adult fly and is involved in immune surveillance mechanisms required for recognition of bacterial and fungal pathogens and regulation of specific innate immune defence gene expression. Activation of dToll receptors in response to infection by specific micro-organisms is thought to require the production of a protein ligand called Spaetzle. The human Toll-like receptors (hTLRs) are also thought to participate in mechanisms of innate immunity and inflammation acting as pattern recognition receptors (PRRs) for bacteria and other micro-organisms. hTLRs are expressed on antigen presenting cells including monocytes and dendritic cells and like dToll play roles in innate immunity. TLRs can elicit pro-inflammatory cytokine production and induce expression of cell surface co-stimulatory receptors required for activation of T-cells. Some hTLRs may help to co-ordinate interactions between cells of the innate and acquired immune systems to orchestrate an integrated immune response to infection.

(13) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 August 2001 (02.08.2001)

(10) International Publication Number
WO 01/55386 A1

(51) International Patent Classification: C12N 15/12, 15/63, 5/10, G01N 33/68, C07K 14/703, 1/628, A61K 31/7083, 39/17, 48/00

(52) International Application Number: PCT/GB01/00299

(53) International Filing Date: 25 January 2001 (25.01.2001)

(54) Filing Language: English

(55) Publication Language: English

(56) Priority Date: 25 January 2000 (25.01.2000) GB 0001704.6

(71) Applicant (for all designated States except US): GLAXO-SMITHKLINE (GB/CB); Corporate Intellectual Property, Two New Horizons court, Brentford, Middlesex TW8 9EP (GB).

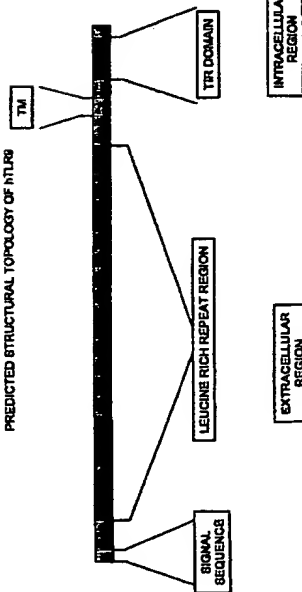
(72) Inventors; and
(73) Investors/Applicants (for US only): LEWIS, Alan, Peter (GB/CB); Glaxo Wellcome plc, Gunnels Wood Road, Stevenage, Hertfordshire SG1 2NY (GB); RAY, Keith, Paul (GB/CB); Glaxo Wellcome plc, Gunnels Wood Road, Stevenage, Hertfordshire SG1 2NY (GB).

Published: with international search report
before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guide to the Preparation of Claims and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) TITLE: TOLL-LIKE RECEPTOR

PREDICTED STRUCTURAL TOPOLOGY OF hTLR9



(57) Abstract: An isolated Toll-like receptor polypeptide comprises the amino acid sequence of SEQ ID NO: 1, a variant or a fragment thereof which has immunomodulatory activity. Polynucleotides encoding such a Toll-like receptor are also described. A method for identification of a substance that modulates Toll-like receptor activity comprises contacting a polypeptide of the invention with a test substance and monitoring for immunomodulatory activity.

2

The specific functions of two mammalian TLRs, TLR2 and TLR4, have recently been identified. TLR2 and TLR4 are involved in mediating host responses to gram positive and gram negative bacteria through recognition of specific bacterial wall components. It has also recently been shown that TLR4 mediates responses to certain viral proteins including respiratory syncytial virus (RSV) (Nature Immunology 1: 398 2000).

Additionally TLRs may form heterodimeric functional complexes. Components of the intracellular signal transduction pathways of some hTLRs appear to be shared with Interleukin-1 (IL-1) receptor transduction pathways. Stimulation of TLR2 and TLR4 leads to activation of NF κ B via an adaptor protein MyD88 and recruitment of the IL-1 receptor associated kinases (IRAKs).

Summary of the invention

A novel Toll-like receptor is now provided which is a screening target for the identification and development of novel pharmaceutical agents which modulate the activity of the receptor and in particular have immunomodulatory activity. These agents may be used in the treatment and/or prophylaxis of inflammatory diseases, cardiovascular diseases, systemic infections and autoimmune diseases, such as asthma, rhinitis, chronic obstructive pulmonary disease (COPD), emphysema, inflammatory bowel disease such as ulcerative colitis and Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, Alzheimer's disease, atherosclerosis, viral, fungal and bacterial infections, septic shock syndrome associated with systemic infection involving gram positive and gram negative bacteria, diabetes, Multiple Sclerosis. These agents may also be used as immunoadjuvants to enhance or alter the immune response in vaccine therapy.

Accordingly, the present invention provides an isolated Toll-like-receptor polypeptide which comprises:

3

- (i) the amino acid sequence of SEQ ID NO: 2;
- (ii) a variant of (i) which has immunomodulatory activity; or
- (iii) a fragment of (i) or (ii) which retains immunomodulatory activity.

Preferably, a variant has at least 80% identity to the amino acid sequence of SEQ ID NO: 2, more preferably at least 95% identity therewith, for example 97% identity therewith.

The invention also provides a polynucleotide encoding a polypeptide of the invention. Such a polynucleotide may be a polynucleotide which encodes a Toll-like receptor polypeptide which has immunomodulatory activity. The polynucleotides of the invention may be DNA or RNA, for example mRNA. A polynucleotide according to the invention comprises:

- (a) the nucleic acid sequence of SEQ ID NO: 1 and/or a sequence complementary thereto;
- (b) a sequence which hybridises under stringent conditions to a sequence as defined in (a);
- (c) a sequence that is degenerate as a result of the genetic code to a sequence as defined in (a) or (b); or
- (d) a sequence having at least 60% percent identity to a sequence as defined in (a), (b) or (c).

The present invention also provides a polypeptide expressed from a polynucleotide according to (a), (b), (c) or (d) above, in particular a polypeptide comprising a toll-like receptor according to the invention, encoded by the mRNA derived from a DNA sequence according to (a) or (b) above, thus the invention provides an isolated toll-like receptor polypeptide which is obtainable by expression *in vitro* or *in vivo* of a DNA molecule comprising the sequence of nucleotides as shown in SEQ ID NO: 1.

The polypeptides of SEQ ID NO 2 and SEQ ID NO 4, which are different isoforms expressed from the multiple exon *tlr9* gene, are herein referred to as TLR9 and TLR9-A, respectively. TLR9-A is encoded by the nucleotide sequence of SEQ ID NO. 3, which is encoded within SEQ ID NO.1 except for the initiating methionine, that is encoded by a second exon as illustrated in figure 1. (see Hemmi et. al. Nature 408, 740-745 2000; Du et al, European Cytokine Network, 11: 362-371, 2000; Chuang and Ulevitch, European Cytokine Network, 11: 372-378 for isolation of the cDNAs, and corresponding sequence database accessions EMBL:AB045180, EMBL:AF259262, EMBL:AF245704).

10

In further aspects of the invention we provide:

- an expression vector capable of expressing a polypeptide of the invention comprising a polynucleotide as defined above.
- a host cell comprising an expression vector of the invention.
- an antibody specific for a polypeptide of the invention.
- a method for identification of a compound that modulates Toll-like receptor activity, comprising contacting a polypeptide of the invention with a test compound and monitoring for immunomodulatory activity.

15

20

Compounds which are identifiable in accordance with this method may be used in the treatment of a subject having a disorder that is responsive to Toll-like receptor modulation such as an inflammatory or cardiovascular disorder or systemic infection or autoimmune disease, including asthma, chronic obstructive pulmonary disease (COPD), emphysema, inflammatory bowel disease such as ulcerative colitis and Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, viral, fungal and bacterial infections, Alzheimer's disease, atherosclerosis, septic shock syndrome associated with systemic infection involving gram positive and gram negative bacteria, diabetes and Multiple Sclerosis. In particular, compounds which are identifiable in accordance with this method may be used in the treatment of a subject having allergic asthma or

30

rhinitis. Further, such compounds may have immunomodulatory activity and be of use in the treatment of, or as adjuvants in vaccination against, bacterial or viral infections or as components of anti-cancer vaccines.

5

Compounds identifiable in accordance with this method include, in particular, synthetic or naturally occurring oligopeptides or polypeptides, oligonucleotides or polynucleotides which bind directly to the Toll-like receptor of the invention, and synthetic or naturally occurring oligopeptides or polypeptides, oligonucleotides or polynucleotides which modulate the Toll-like receptor of the present invention via one or more intermediate signal transducers. Such oligo- or polynucleotides may be "CG-rich" sequences or sequences including one or more unmethylated CpG nucleotide pairs.

10

In an alternative aspect of the invention, a polypeptide comprises a fragment or variant of SEQ ID NO 2 which is capable of inhibiting the activity of TLR9 or TLR9-A, for use in the treatment of an immune or inflammatory disorder.

15

In a further aspect of the invention, a polypeptide or polynucleotide in accordance with the invention or a compound identifiable in accordance with the invention is provided for use as an adjuvant or as an immunotherapeutic agent, for example in a vaccine.

20

Brief Description of the Sequences

SEQ ID NO: 1 is the amino acid sequence of human protein TLR9 and its encoding DNA;

25

SEQ ID NO: 2 is the amino acid sequence alone of TLR9;

SEQ ID NO: 3 is the amino acid sequence of human protein TLR9-A and its encoding cDNA (EMBL:AF259262);

SEQ ID NO: 4 is the amino acid sequence alone of TLR9-A (Hemmi et al.).

30

Brief Description of the Drawings

Fig. 1 is a diagrammatic illustration showing the exon arrangement encoding TLR9 and TLR9-A;

Fig. 2 shows tissue distribution data for TLR9 (using a human tissue cDNA plate). The profile shows predominant expression in tonsil and adenoid tissues with lower levels of expression detected in adipose, adrenal, foetal brain, cerebellum, jejunum, lung, myometrium, omentum, head of pancreas, rectum, skeletal muscle, spleen and thymus tissues;

Fig. 3 shows tissue distribution data for TLR9 (using a human disease cDNA plate). The profile shows predominant expression in lung tissue, bone marrow and PBMC with lower levels of expression detected in some colon, breast and brain/cerebellum samples;

Fig. 4 illustrates, in diagrammatic form, the predicted structural topology of human TLR9 - "TM" is the transmembrane portion, "TIR" is the cytosolic region conserved among Interleukin and toll-like receptors known as the Toll Interleukin Receptor domain.

Detailed Description of the Invention

A single open reading frame was identified in genomic DNA (SEQ ID No.1), which encodes a protein of 1055 residues, predicted molecular weight 118,515 (PeptideSort - GCG Software) and the amino acid sequence shown in SEQ ID No.2. This sequence included the TIR domain common to Toll-like receptors and members of the interleukin-1 receptor family e.g. IL1RI, and the N-terminal sequence contains structural features as shown in figure 4. These features include, in order from the N- to the C-terminus, a predicted signal sequence with a potential cleavage site between residues 48 and 49 or 50 and 51 (SPScan in GCG; SignalP), a leucine-rich repeat motif domain, a potential transmembrane region and the Toll/IL-1R homologous region (TIR; Pfam Database). These motifs confirmed that the protein was likely to be expressed as a single transmembrane receptor-like molecule belonging to the TLR rather than the IL-1R

family and therefore, it was designated TLR9 based on existing published and in-house nomenclature.

Throughout the present specification and the accompanying claims the words "comprise" and "include" and variations such as "comprises", "comprising", "includes" and "including" are to be interpreted inclusively. That is, these words are intended to convey the possible inclusion of other elements or integers not specifically recited, where the context allows.

The present invention relates to a human Toll-like receptor, referred to herein as TLR9, and variants or fragments thereof. Sequence information for TLR9 is provided in SEQ ID NO: 1 (nucleotide and amino acid) and in SEQ ID NO: 2 (amino acid only). A polypeptide of the invention consists essentially of the amino acid sequence of SEQ ID NO: 2 or of a functional variant of that sequence. One important variant of TLR9 is TLR9-A, sequence information for which is provided in SEQ ID NO: 3 (nucleotide and amino acid) and in SEQ ID NO: 4 (amino acid only).

The polypeptides are provided in isolated form. The term "isolated" is intended to convey that the polypeptide is not in its native state, insofar as it has been purified at least to some extent or has been synthetically produced, for example by recombinant methods. The polypeptide may be mixed with carriers or diluents which will not interfere with the intended purpose of the polypeptide and still be regarded as substantially isolated. The term "isolated" therefore includes the possibility of the polypeptide being in combination with other biological or non-biological material, such as cells, suspensions of cells or cell fragments, proteins, peptides, expression vectors, organic or inorganic solvents, or other materials where appropriate, but excludes the situation where the polypeptide is in a state as found in nature.

A polypeptide of the invention may also be in a substantially purified form, in which case it will generally comprise the polypeptide in a preparation in which more than 50%, e.g. more than 80%, 90%, 95% or 99%, by weight of the polypeptide in the preparation is a polypeptide of the invention. Routine methods, can be employed to purify and/or synthesise the proteins according to the invention. Such methods are well understood by persons skilled in the art, and include techniques such as those disclosed in Sambrook *et al.*, Molecular Cloning: a Laboratory Manual, 2nd Edition, CSH Laboratory Press (1989), the disclosure of which is included herein in its entirety by way of reference.

The term "variants" refers to polypeptides which have the same essential character or basic biological functionality as TLR9. The essential character of TLR9 can be defined as that of a Toll-like receptor. In particular, it refers to a polypeptide which has an immunomodulatory activity. In one aspect of the invention, a polypeptide of the invention may activate NFkB or may elicit pro-inflammatory cytokine production or induce expression of cell surface co-stimulatory receptors required for activation of T-cells.

Alternatively, a variant of the polypeptide of the invention is one which exhibits binding to the same ligand as TLR9. Such ligand binding may be assayed using the assays described below.

In other aspects of the invention a variant is one which does not show the same function as TLR9 but which may be used to inhibit this function. For example, a variant polypeptide for use in an assay or therapy is one which inhibits TLR9 activity, for example by inhibiting or compelling out ligand binding or receptor complex formation by TLR9. Alternatively, a variant may be one which inhibits ligand binding to TLR9. Such a variant may inhibit activation of NFkB or inhibit cytokine production and expression of cell surface co-stimulatory receptors.

Such inhibitors may be used as immunomodulators to reduce inappropriate TLR activation in asthma or other chronic inflammatory diseases, or septic shock.

To determine whether a variant has the same essential function as TLR9, the immunomodulatory activity can be determined by monitoring the effect of a substance on different immune responses. For example the effect of the substance under test on NFkB activation mediated through binding the polypeptide of the present invention may be monitored. This can be carried out, for example, by co-transfection of a construct expressing the polypeptide with a construct containing a reporter gene, such as secreted placental alkaline phosphatase, under the control of a suitable NFkB-responsive promoter and monitoring for expression of the reporter gene.

Alternatively, other immunomodulatory activity such as the production of cytokines can be determined by monitoring cytokine production following incubation of a test substance with a cell expressing a polypeptide of the invention. Such assays may be carried out in the presence or absence of additional T-lymphocytes to assess the effect of such cytokines, or the direct action of a polypeptide of the invention, on such T-lymphocytes to thus determine immunomodulatory activity.

Alternatively, the Toll-like receptor functionality is as a peptide which binds a ligand of TLR9, inhibits immunomodulatory activity by TLR9 or inhibits ligand binding to TLR9 and can be determined by an assay as described below.

Preferably, a polypeptide of the invention will show the structural features associated with a Toll-like receptor. Preferably, a polypeptide of the invention, or a functional fragment thereof, contains one or more of the following structural features associated with a Toll-like receptor: an extracellular region containing leucine-rich repeat motif and cysteine-rich regions involved in ligand binding; a

single hydrophobic transmembrane region; and a C-terminal cytoplasmic domain sharing homology with other TLRs and members of the IL-1 receptor family.

Typically, polypeptides with more than about 65% identity, preferably at least 80% or at least 90% and particularly preferably at least 95%, at least 97%, or at least 99% identity, with the amino acid sequences of SEQ ID NO: 2 over a region of at least 20, preferably at least 30, at least 40, at least 60 or at least 100 contiguous amino acids or over the full length of SEQ ID NO: 2, are considered as variants of the proteins. Identity is calculated using the widely used GCG (University of Wisconsin) suite of programs and preferably using the distances software (correction method). Such variants may include allelic variants and the deletion, modification or addition of single amino acids or groups of amino acids within the protein sequence, as long as the peptide maintains the basic biological functionality of the Toll-like receptor, having a similar function to TLR9 or inhibits such function such as preventing ligand binding or TLR9 mediated activation. Such variants also include isoforms such as TLR9-A, which is 23 amino acids (or 2.2%) shorter than TLR9 (see SEQ ID NO: 2 and SEQ ID NO: 4) and thus shows 97.8 identity therewith. Transcription of the nucleotide sequence presented in SEQ ID NO: 1 can result, due to variable mRNA splicing involving a second exon encoding an alternative initiating methionine, in an mRNA having the sequence of SEQ ID NO: 3 which, when translated, results in the polypeptide TLR9-A depicted in SEQ ID NO: 4.

Amino acid substitutions may be made, for example from 1, 2 or 3 to 10, 20 or 30 substitutions. The modified polypeptide generally retains activity as a TLR9 receptor or inhibitor of TLR9 receptor activity. Conservative substitutions may be made, for example according to the following Table. Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other.

ALIPHATIC	Non-polar	GAP
		ILV
	Polar-uncharged	CSM
		NQ
	Polar-charged	DE
KR		
AROMATIC		HFWY

Shorter polypeptide sequences are within the scope of the invention. For example, a peptide of at least 20 amino acids or up to 50, 60, 70, 80, 100 or 150 amino acids in length is considered to fall within the scope of the invention as long as it demonstrates the basic biological functionality of TLR9 or inhibits TLR9. In accordance with this aspect of the invention the peptide may also comprise an epitope of TLR9 for generation of antibodies. In particular, but not exclusively, this aspect of the invention encompasses the situation when the protein is a fragment of the complete protein sequence and may represent a ligand-binding region (N-terminal extracellular domain) or an effector binding region (C-terminal intracellular domain). Fragments from which the C-terminus has been removed may be used as decoy receptors. Other fragments such as a secreted or soluble form of the receptor may be generated for use in an assay or in therapy in accordance with the invention. Such fragments can also be used to raise anti-TLR9 antibodies.

Polypeptides of the invention may be chemically modified, e.g. post-translationally modified. For example, they may be glycosylated or may comprise modified amino acid residues. They may also be modified by the addition of histidine residues or an epitope tag for example by a (His) 8 or (His) 6 sequence or a HA, T7, Myc or Flag tag to assist their purification or detection. They may be modified by the addition of a signal sequence to promote insertion into the

cell membrane. Such modified polypeptides fall within the scope of the term "polypeptide" of the invention.

The invention also includes nucleotide sequences that encode for TLR9 or variants thereof as well as nucleotide sequences which are complementary thereto. The nucleotide sequence may be RNA or DNA including genomic DNA, synthetic DNA or cDNA. Preferably the nucleotide sequence is a DNA sequence and most preferably, a cDNA sequence. Nucleotide sequence information is provided in SEQ ID NO: 1. Such nucleotides can be isolated from human cells or synthesised according to methods well known in the art, as described by way of example in Sambrook *et al.* Such nucleotides can typically be isolated from activated cells of the immune system, heart, lung, pancreatic islet cells and lymph nodes, adenoid and tonsil tissues. Figures 2 and 3 show the tissue distribution of RNA encoding TLR9, as determined by extraction of total RNA from normal or disease tissue or cells which is then used to generate cDNA for real time quantitative PCR using suitable primers and probes (TaqMan analysis) to assess expression patterns. The profiles show differential expression across tissues tested and predominance to sites containing inflammatory cell types.

Typically a polynucleotide of the invention comprises a contiguous sequence of nucleotides which is capable of hybridising under selective conditions to the coding sequence or the complement of the coding sequence of SEQ ID NO: 1.

A polynucleotide of the invention can hybridize to the coding sequence or the complement of the coding sequence of SEQ ID NO: 1 (or of SEQ ID NO: 3) at a level significantly above background. Background hybridisation may occur, for example, because of other cDNAs present in a cDNA library. The signal level generated by the interaction between a polynucleotide of the invention and the coding sequence or complement of the coding sequence of SEQ ID NO: 1 or of SEQ ID NO: 3 is typically at least 10 fold, preferably at least 100 fold, as intense

as interactions between other polynucleotides and the coding sequence of SEQ ID NO: 1 or of SEQ ID NO: 3. The intensity of interaction may be measured, for example, by radiolabelling the probe, e.g. with ^{32}P . Selective hybridisation may typically be achieved using conditions of low stringency (0.03M sodium chloride and 0.03M sodium citrate at about 40°C), medium stringency (for example, 0.03M sodium chloride and 0.03M sodium citrate at about 50°C) or high stringency (for example, 0.03M sodium chloride and 0.03M sodium citrate at about 60°C).

The coding sequence of SEQ ID NO: 1 may be modified by nucleotide substitutions, for example from 1, 2 or 3 to 10, 25, 50 or 100 substitutions. The polynucleotides of the present invention may alternatively or additionally be modified by one or more insertions and/or deletions and/or by an extension at either or both ends. The modified polynucleotide generally encodes a polypeptide which has Toll-like receptor activity or inhibits the activity of TLR9. Degenerate substitutions may be made and/or substitutions may be made which would result in a conservative amino acid substitution when the modified sequence is translated, for example as shown in the Table above.

A nucleotide sequence of the invention which is capable of selectively hybridising to the complement of the DNA coding sequence of SEQ ID NO: 1 will generally have at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98% or at least 99% sequence identity to the coding sequence of SEQ ID NO: 1 over a region of at least 20, preferably at least 30, for instance at least 40, at least 60, more preferably at least 100 contiguous nucleotides or most preferably over the full length of SEQ ID NO: 1. Methods of measuring nucleic acid and protein homology are well known in the art. For example the UWGCG Package provides the BESTFIT program which can be used to calculate homology (Devereux *et al* 1984). Similarly the PILEUP and BLAST algorithms can be used to line up sequences (for example as described in Altschul 1993,

and Aitschui *et al* (1990). Many different settings are possible for such programs. In accordance with the invention, the default settings may be used.

Any combination of the above mentioned degrees of sequence identity and minimum sizes may be used to define polynucleotides of the invention, with the more stringent combinations (i.e. higher sequence identity over longer lengths) being preferred. Thus, for example a polynucleotide which has at least 80% sequence identity over 25, preferably over 30 nucleotides forms one aspect of the invention, as does a polynucleotide which has at least 95% sequence identity over 40 nucleotides. The most preferred sequences have at least 70% sequence identity over at least 70% of the full length of the sequence provided by SEQ ID NO: 1.

The nucleotides according to the invention have utility in production of the proteins according to the invention, which may take place *in vitro*, *in vivo* or *ex vivo*. The nucleotides may be involved in recombinant protein synthesis or indeed as therapeutic agents in their own right, utilised in gene therapy techniques. Nucleotides complementary to those encoding TLR9, or antisense sequences, may also be used in gene therapy, such as in strategies for down regulation of expression of the proteins of the invention.

Polynucleotides of the invention may be used as a primer, e.g. a PCR primer, a primer for an alternative amplification reaction, a probe e.g. labelled with a revealing label by conventional means using radioactive or non-radioactive labels, or the polynucleotides may be cloned into vectors.

Such primers, probes and other fragments will preferably be at least 10, preferably at least 15 or at least 20, for example at least 25, at least 30 or at least 40 nucleotides in length. They will typically be up to 40, 50, 60, 70, 100 or 150 nucleotides in length. Probes and fragments can be longer than 150

nucleotides in length, for example up to 200, 300, 400, 500 nucleotides in length, or even up to a few nucleotides, such as five or ten nucleotides, short of the coding sequence of SEQ ID NO: 1.

The present invention also includes expression vectors that comprise nucleotide sequences encoding the proteins or variants thereof of the invention. Such expression vectors are routinely constructed in the art of molecular biology and may for example involve the use of plasmid DNA and appropriate initiators, promoters, enhancers and other elements, such as for example polyadenylation signals which may be necessary, and which are positioned in the correct orientation, in order to allow for protein expression. Other suitable vectors would be apparent to a person skilled in the art. By way of further example in this regard we refer to Sambrook *et al*.

Polynucleotides according to the invention may also be inserted into the vectors described above in an antisense orientation in order to provide for the production of antisense RNA. Antisense RNA or other antisense polynucleotides may also be produced by synthetic means. Such antisense polynucleotides may be used as test compounds in the assays of the invention or may be useful in a method of treatment of the human or animal body by therapy.

Preferably, a polynucleotide of the invention or for use in the invention in a vector is operably linked to a control sequence which is capable of providing for the expression of the coding sequence by the host cell, i.e. the vector is an expression vector. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A regulatory sequence, such as a promoter, "operably linked" to a coding sequence is positioned in such a way that expression of the coding sequence is achieved under conditions compatible with the regulatory sequence.

The vectors may be for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of the said polynucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial plasmid or a resistance gene for a fungal vector. Vectors may be used *in vitro*, for example for the production of DNA or RNA or used to transfect or transform a host cell, for example, a mammalian host cell. The vectors may also be adapted to be used *in vivo*, for example in a method of gene therapy.

Promoters and other expression regulation signals may be selected to be compatible with the host cell for which expression is designed. For example, yeast promoters include *S. cerevisiae* GAL4 and ADH promoters, *S. pombe* *nmt1* and *adh* promoter. Mammalian promoters include the metallothionein promoter which can be induced in response to heavy metals such as cadmium. Viral promoters such as the SV40 large T antigen promoter or adenovirus promoters may also be used. All these promoters are readily available in the art.

20 Mammalian promoters, such as β -actin promoters, may be used. Tissue-specific promoters may be used. Viral promoters may also be used, for example the Moloney murine leukaemia virus long terminal repeat (MMLV LTR), the rous sarcoma virus (RSV) LTR promoter, the SV40 promoter, the human cytomegalovirus (CMV) IE promoter, adenovirus, HSV promoters (such as the HSV IE promoters), or HPV promoters, particularly the HPV upstream regulatory region (URR). Viral promoters are readily available in the art.

25

The vector may further include sequences flanking the polynucleotide which comprise sequences homologous to eukaryotic genomic sequences, preferably mammalian genomic sequences, or viral genomic sequences. This will allow the

Introduction of the polynucleotides of the invention into the genome of eukaryotic cells or viruses by homologous recombination. In particular, a plasmid vector comprising the expression cassette flanked by viral sequences can be used to prepare a viral vector suitable for delivering the polynucleotides of the invention to a mammalian cell. Other examples of suitable viral vectors include herpes simplex viral vectors and retroviruses, including lentiviruses, adenoviruses, adeno-associated viruses and HPV viruses (such as HPV-16 or HPV-18). Gene transfer techniques using these viruses are known to those skilled in the art. Retrovirus vectors for example may be used to stably integrate the polynucleotide giving rise to the RNA into the host genome. Replication-defective adenovirus vectors by contrast remain episomal and therefore allow transient expression.

The invention also includes cells that have been modified to express the TLR9 polypeptide or a variant thereof. Such cells include transient, or preferably stable higher eukaryotic cell lines, such as mammalian cells or insect cells, lower eukaryotic cells, such as yeast or prokaryotic cells such as bacterial cells. Particular examples of cells which may be modified by insertion of vectors encoding for a polypeptide according to the invention include mammalian HEK293T, CHO, HeLa and COS cells. Preferably the cell line selected will be one which is not only stable, but also allows for mature glycosylation and cell surface expression of a polypeptide. Cells such as T-cells, monocytes or dendritic cells expressing the receptor may be used for example in screening. Expression may be achieved in transformed oocytes. A polypeptide of the invention may be expressed in cells of a transgenic non-human animal, preferably a mouse. A transgenic non-human animal expressing a polypeptide of the invention is included within the scope of the invention.

It is also possible for the proteins of the invention to be transiently expressed in a cell line or on a membrane, such as for example in a baculovirus expression

18

system. Such systems, which are adapted to express the proteins according to the invention, are also included within the scope of the present invention.

5 According to another aspect, the present invention also relates to antibodies (either polyclonal or preferably monoclonal antibodies, chimeric, single chain, Fab fragments) which have been raised by standard techniques and are specific for a polypeptide of the invention. Such antibodies could for example, be useful in purification, isolation or screening methods involving immunoprecipitation techniques and may be used as tools to further elucidate the function of TLR9 or a variant thereof, or indeed as therapeutic agents in their own right. Antibodies may also be raised against specific epitopes of the proteins according to the invention. Such antibodies may be used to block ligand binding to the receptor. Alternatively an antibody may be provided which acts as an agonist, to cross link receptors of the invention to mediate receptor activity. An antibody, or other compound, "specifically binds" to a protein when it binds with high affinity to the protein for which it is specific but does not bind or binds with only low affinity to other proteins. A variety of protocols for competitive binding or immunoradiometric assays to determine the specific binding capability of an antibody are well known in the art (see for example Maddox et al 1993). Such immunosays typically involve the formation of complexes between the "specific protein" and its antibody and the measurement of complex formation.

20 An important aspect of the present invention is the use of polypeptides according to the invention in screening methods to identify compounds that may act as modulators of Toll-like receptor activity. Any suitable form may be used for the assay to identify a modulator of TLR9 activity. In general terms, such screening methods may involve contacting a polypeptide of the invention with a test compound and then measuring receptor activity.

30

19

Screening methods may alternatively involve contacting a polypeptide of the invention with a test compound and then monitoring for the effect on immunomodulatory activity.

5 The binding of the substance to a polypeptide in the invention can be determined directly. For example, a radiolabelled test substance can be incubated with a polypeptide of the invention and so that binding of the test substance to the polypeptide can be monitored. Typically, the radiolabelled test substance can be incubated with cell membranes containing the polypeptide until equilibrium is reached. The membranes can then be separated from a non-bound test substance and dissolved in scintillation fluid to allow the radioactive content to be determined by scintillation counting. Non-specific binding of the test substance may also be determined by repeating the experiments in the presence of a saturating concentration of a non-radioactive ligand. Preferably, a binding curve is constructed by repeating the experiment with various concentrations of the test substance.

15 Cell based assays may also be carried out, for example using a cell expressing the TLR9 receptor, and contacting the cell with another cell to look for ligand binding or activation of TLR9-mediated pathways such as NFkB activation. Alternatively cells expressing TLR9 constitutively may be provided for use in assays for TLR9 function. Such constitutively expressed TLR9 may demonstrate TLR9 activity in the absence of ligand binding. Additional test substances may be introduced in any assay to look for inhibitors of ligand binding or inhibitors of TLR9-mediated activity. Assays are preferably carried out using cells expressing TLR9, and incubating such cells with the test substance optionally in the presence of TLR9 ligand. Alternatively an antibody may be used to complex TLR9 and thus mediate TLR9-activity. Test substances may then be added to assess the effect on such activity.

In preferred aspects, a host cell is provided expressing the receptor, or a receptor complex of TLR8 (or TLR8-A) comprising a homodimer, a heterodimer of TLR8 (or TLR8-A) with another Toll-like receptor, or a complex of TLR9 (or TLR9-A) with protein cofactors, and containing an NFkB responsive reporter construct. The host cell is treated with a substance under test for a defined time. The expression of the reporter gene, such as secreted placental (SP) alkaline phosphatase or luciferase is assayed. The assay enables determination of whether the addition of compounds inhibits the induction of the response in target cells. Alternatively the assay may be carried out to identify cytokine production or it may be carried out in the presence of T-cells to identify inducement of co-stimulatory receptors required for activation of T-cells.

Assays may also be carried out to identify modulators of receptor shedding. A polypeptide of the invention can be cleaved from the cell surface. Shedding the receptor would act to down regulate receptor signalling. Thus, cell based assays may be used to screen for compounds which promote or inhibit receptor-shedding. Assays may also be carried out to identify substances which modify TLR8 receptor expression for example substances which down regulate expression. Such assays may be carried out for example by using antibodies for TLR8 to monitor levels of TLR8 expression.

Additional control experiments may be carried out. Assays may also be carried out using known ligands of other Toll-like receptors to identify ligands which are specific for polypeptides of the invention. Preferably, the assays of the invention are carried out under conditions which would result in Immunomodulatory NFkB mediated activity in the absence of the test substance, to identify inhibitors of Toll-like receptor mediated activity, or agents which inhibit ligand-induced Toll-like receptor activity.

Suitable test substances which can be tested in the above assays include combinatorial libraries, defined chemical entities, peptides and peptide mimetics, oligonucleotides and natural product libraries, such as display (e.g. phage display libraries) and antibody products. In a preferred embodiment, the test substance is a variant peptide of the invention. In particularly preferred embodiments, suitable test substances which may be candidate ligands for binding to and modulation of TLR8 or TLR8-A include synthetic or naturally occurring oligonucleotides or polynucleotides which bind directly to the Toll-like receptor or which modulate the Toll-like receptor of the present invention via one or more intermediate signal transducers. Such oligo- or polynucleotides may be "CG-rich" sequences or sequences including one or more unmethylated CpG nucleotide pairs.

The assay may be carried out using full length TLR8 to identify a variant peptide which interferes with TLR8 mediated activity, for example by inhibiting ligand binding.

Test substances may be used in an initial screen of, for example, 10 substances per reaction, and the substances of these batches which show inhibition or activation tested individually. Test substances may be used at a concentration of from 1nM to 1000µM, preferably from 1µM to 100µM, more preferably from 1µM to 10µM.

A protein-binding assay may be developed using a polypeptide of the invention, preferably one encoding the extracellular ligand-binding domain, to identify novel protein ligands of TLR8. Particular examples may be screening of a human cDNA expression library for protein ligands of TLR8 by yeast 2-hybrid protein interaction assay (e.g. as described in International Patent Application No. WO98/49294).

22

Another aspect of the present invention is the use of polynucleotides encoding the TLR9 polypeptides of the invention to identify mutations in TLR9 genes which may be implicated in human disorders or to identify cells in which TLR9 is expressed. Identification of such mutations may be used to assist in diagnosis of immune system, lung, kidney, heart or other disorders or susceptibility to such disorders and in assessing the physiology of such disorders. In particular the polynucleotides of the invention may assist in diagnosis of asthma and rheumatoid arthritis. For example, a SNP (single nucleotide polymorphism) has been identified in the genomic DNA encoding TLR9 (G/A nucleotide: The SNP Consortium database accession number TSC0164834). This single base pair change lies in the DNA region encoding the 23 N-terminal residues of TLR9, and this region is spliced out of the mRNA encoding TLR9-A. The nucleotide at this SNP position may affect the efficiency of mRNA splicing in the two different variants - a G at this position may possibly disrupt the splicing machinery and an A might lead to more efficient splicing. Additionally, the presence of a G as compared to an A in an unspliced mRNA would introduce a stop codon and result in different N terminal protein sequences upon translation of that mRNA, thus the two polymorphic variants of the *tlr9* gene may encode receptors which have differing expression levels and/or differing functional activity levels. The present invention provides a diagnostic tool for determining the polymorphic variant in an individual by detecting the DNA sequence at the SNP site. Such a tool may incorporate a nucleotide probe specific for one or other of the polymorphic variants, for example an oligonucleotide of from 5 to 50, preferably 5-20 nucleotides, complementary to a fragment of the nucleotide sequence of SEQ ID No. 1 which extends over the SNP site or a fragment complementary to that sequence with the exception of the single nucleotide change (G to A) at the SNP site. The present invention also provides a method of detecting a polymorphic variant in the *tlr9* gene by determining the sequence of nucleotides at and around the SNP site identified by the SNP consortium database accession number

23

TSC0164834, in particular by determining whether the nucleotide at that SNP site is a G or an A.

Another aspect of the present invention is the use of the compounds that have been identified by screening techniques referred to above in the treatment or prophylaxis of disorders which are responsive to regulation of TLR9 receptor activity. In addition, variant peptides of the invention which inhibit TLR9-mediated activity, for example which inhibit ligand binding or prevent hTLR9 immunomodulatory activity may be used in the treatment or prophylaxis of such disorders. Antibodies which recognise TLR9 may similarly be used in therapy.

In particular, such compounds may be used in the treatment of inflammatory, cardiovascular, systemic infection or autoimmune disease. The compounds may be used to treat bacterial, viral or fungal infections, asthma, rhinitis, chronic obstructive pulmonary disease (COPD), emphysema, an inflammatory bowel disease such as ulcerative colitis and Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, Alzheimer's disease, atherosclerosis, septic shock syndrome associated with systemic infection involving gram positive and gram negative bacteria, diabetes, Multiple Sclerosis.

In an alternative aspect, the invention provides agents which activate TLR9 mediated immunomodulation for use as an immunoadjuvant, or TLR9, and variants thereof, or polynucleotides or DNA encoding a polypeptide of the invention which may be administered for use as immunoadjuvants to enhance or alter the immune response in an individual to an antigen.

The compounds identified according to the screening methods outlined above may be formulated with standard pharmaceutically acceptable carriers and/or excipients as is routine in the pharmaceutical art, and as fully described in Remington's Pharmaceutical Sciences, Mack Publishing Company, Eastern

Pennsylvania 17th Ed. 1985, the disclosure of which is included herein of its entirety by way of reference.

5 The compounds may be administered by enteral or parenteral routes such as via oral, buccal, anal, pulmonary, intravenous, intra-arterial, intramuscular, intraperitoneal, topical or other appropriate administration routes. The dose of a compound to be administered may be determined according to various parameters, especially according to the substance used; the age, weight and condition of the patient to be treated; the route of administration; and the required regimen. Again, a physician will be able to determine the required route of administration and dosage for any particular patient. A typical daily dose is from about 0.1 to 50 mg per kg of body weight, according to the activity of the compound, the age, weight and conditions of the subject to be treated, the type and severity of the disease and the frequency and route of administration. Preferably, daily dosage levels are from 5 mg to 2 g.

15 Nucleic acid encoding an inhibitor of TLR9 activity may be administered to the mammal. In an alternative aspect of the invention, nucleic acid encoding TLR9 or a variant thereof may be administered to the animal. Such a variant shows immunomodulatory activity of TLR9 such as inducing cytokine production and expression of cell surface co-stimulatory receptors required for activation of T-cells. Nucleic acid, such as RNA or DNA, and preferably, DNA, is provided in the form of a vector, such as the polynucleotides described above, which may be expressed in the cells of the mammal.

25 Nucleic acid encoding the peptide may be administered to the animal by any available technique. For example, the nucleic acid may be introduced by injection, preferably intradermally, subcutaneously or intramuscularly. Alternatively, the nucleic acid may be delivered directly across the skin using a nucleic acid delivery device such as particle-mediated gene delivery. The

nucleic acid may be administered topically to the skin, or to the mucosal surfaces for example by intranasal, oral, intravaginal, intrarectal administration.

5 Uptake of nucleic acid constructs may be enhanced by several known transfection techniques, for example those including the use of transfection agents. Examples of these agents includes cationic agents, for example, calcium phosphate and DEAE-Dextran and lipofectants, for example, lipofectam and transfectam. The dosage of the nucleic acid to be administered can be altered. Typically the nucleic acid is administered in the range of 1pg to 1mg, preferably to 1pg to 10µg nucleic acid for particle mediated gene delivery and 10µg to 1mg for other routes.

Examples

15 Example 1 - Screening for compounds which exhibit protein modulating activity

Mammalian cells, such as HeK293, CHO and HeLa cells over-expressing TLR9 protein are generated for use in the assay. 96 and 384 well plate, high throughput screens (HTS) are employed. TLR9 induced cell activation can be monitored through activation of a signal transduction event (typically resulting in activation of NFκB or AP-1) or transcriptional activation of a reporter gene (typically regulated via NFκB or AP-1 responsive elements). TLR9 induced activation of a reporter gene such alkaline phosphatase or luciferase can be easily assessed using a suitable colourimetric or fluorimetric assay to measuring production. Such assays may be used to identify receptor antagonists capable of blocking ligand induced TLR9 activation, inhibitors capable of blocking TLR9 intracellular signal transduction or receptor agonists capable of activating TLR9. Secondary screening involves evaluation of TLR9 modulators in disease related

28

cells. Tertiary screens involve the study of modulators in rat and mouse models of disease relevant to the target.

27

SEQUENCE LISTING

5 <110> GLAXO GROUP LIMITED
<120> PROTEIN
<130> P78721 GCM SER
10 <140>
<141>
<150> 2
<170> Patentin Ver. 2.1
15 <210> 1
<211> 3165
<212> DNA
<213> Homo sapiens
20 <220>
<221> CDS
<222> (1)..(3165)
25 <400> 1
atg ccc atg aag tgg agt ggg tgg agt tgg agc tgg ggg ccg gcc act 48
Met Pro Met Lys Trp Ser Gly Trp Arg Trp Ser Trp Gly Pro Ala Thr 10
1 5 10 15
cac aca gcc ctc cca ccc cca cag ggt ttc tgc cgc agc gcc ctg cac 96
30 His Thr Ala Leu Pro Pro Gln Gly Phe Cys Arg Ser Ala Leu His 25 30
cgg ctg tct ctc ctg gtc cag gcc atc atg ctg gcc atg acc ctg gcc 144
Pro Leu Ser Leu Leu Val Gln Ala Ile Met Leu Ala Met Thr Leu Ala 35 40 45
35 ctg ggt acc ttg cct gcc ttc cta ccc tgt gag ctc cag ccc cac gcc 192
Leu Gly Thr Leu Pro Ala Phe Leu Pro Cys Glu Leu Gln Pro His Gly 50 55 60
ctg gtc aac tgc aac tgg ctg ttc ctg aag tct gtc ccc cac ttc tcc 240
Leu Val Asn Cys Asn Trp Leu Phe Leu Lys Ser Val Pro His Phe Ser 65 70 75 80
40 atg gca gca ccc cgt gcc aat gtc acc agc ctt tcc ttg tcc tcc aac 288
Met Ala Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn 85 90 95
cgc atc cac cac ctc cat ggt tct gac ttt gcc cac ctg ccc agc ctg 336
45 Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu 100 105 110
cgg cat ctc aac ctc aag tgg aac tgc ccg cgt gtt gcc ctc agc ccc 384
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Pro Val Gly Leu Ser Pro 115 120 125
50 atg cac ttc ccc tgc cac atg acc atc gag ccc agc acc ttc ttg ggt 432
Met His Phe Pro Cys His Met Thr Ile Glu Pro Ser Thr Phe Leu Ala 130 135 140

480 gtr ccc acc ctg gaa gag cta aac cag ggc tac aac atc atg act
 Val Pro Thr Leu Glu Glu Leu Asn Leu Ser Tyr Asn Asn Ile Met Thr
 145 150 155 160
 528 gtr cct ggc ctg ccc aaa tcc ctc cta ctg tcc ctc agc act acc
 Ile Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn
 435 440 445
 576 aac atc ctg atg cta gac tct gcc atc ctc ggc ctg cat gcc ctg
 Asn Ile Met Leu Asp Ser Ala Ser Leu Ala Gly Leu His Ala Leu
 180 185 190
 624 cgc ttc cta ttc atg gac ggc aac tct tat tac aag aac ccc tgc agt
 Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Tyr Lys Asn Pro Cys Arg
 195 200 205
 672 cag gaa ctg gag ggc ccc ggt gcc ctc ctc ggc ctg ggc aac ctc
 Gln Ala Leu Glu Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu
 210 215 220
 720 aac cag ctg cta ctc aag tac aac aac ctc act gtr gtr ccc ggc aac
 Thr His Leu Ser Leu Lys Tyr Asn Asn Leu Thr Val Pro Arg Asn
 225 230 235 240
 768 ctg cct tcc agc ctg gaa tat ctg ctg tta tta tta aac cgc atc gtc
 Leu Pro Ser Ser Leu Glu Tyr Leu Leu Leu Ser Tyr Asn Arg Ile Val
 245 250 255
 816 aac ctg gtr cct ggc gac ctg ggc act ctg acc ggc ctg gtr ctc
 Lys Leu Ala Pro Glu Asp Leu Ala Asn Thr Thr Ala Leu Arg Val Leu
 260 265 270
 864 gtr ctg ggc gaa atc tgc ggc ctg tgc gac gac gct ccc aac ccc tgc
 Asp Val Gly Asn Cys Arg Arg Cys Asp His Ala Pro Asn Pro Cys
 275 280 285
 912 atg gtr tgc cct gtr ccc ttc ccc cag cta ctc ccc gac acc ttc agc
 Met Glu Cys Pro Arg His Phe Pro Glu Leu His Pro Asp Thr Phe Ser
 290 295 300
 960 cgc ctg ggc ctt gaa ggc ctg gtr ang gac agt tct tct ttc
 His Leu Ser Arg Leu Glu Gly Leu Val Leu Lys Asp Ser Leu Ser
 305 310 315 320
 1008 tgc ctg atc ggc atg tgc ttc gtr gtr gtr gtr aac ctc cga gtr ctg
 Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu
 325 330 335
 1056 gac ctg agt gaa aac ttc ctc tac aaa tgc atc act aaa acc aag ggc
 Asp Leu Ser Glu Asn Phe Leu Tyr Lys Cys Ile Thr Lys Thr Lys Ala
 340 345 350
 1104 ttc cag ggc cta aca cag ctg ggc agt ctc atc ctg ttc ttc aat tac
 Phe Gln Gly Leu Thr Gln Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr
 355 360 365
 1152 cca aag agt gtr ttc ttc ggc ctc ctg ttc ctg ggc cct ttc ttc gtr
 Gln Lys Arg Val Ser Phe Ala His Leu Ser Leu Leu Ala Pro Ser Phe Gly
 370 375 380
 1200 ggc ctg gtr ggc ctg aag gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Ser Leu Val Ala Leu Lys Glu Leu Asp Met His Gly Ile Phe Phe Arg
 385 390 395 400
 1248 tca ctc gtr ggc acc agt ctc cgc ctg gtr gtr gtr gtr gtr gtr gtr
 50 Ser Leu Asp Glu Thr Thr Leu Arg Pro Leu Ala Arg Leu Pro Met Leu
 405 410 415
 1296 cag act ctg gtr ctg cag atg aac ttc atc aac cag ggc ctc gtr

Gln Thr Leu Arg Leu Gln Met Asn Phe Ile Asn Gln Ala Gln Leu Gly
 420 425 430
 1344 atc ttc agt gcc ttc cct ggc ctg gtr gtr gtr gtr gtr gtr gtr gtr
 Ile Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn
 435 440 445
 1392 cgc atc agc gaa gct tgc gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Arg Ile Ser Gly Ala Ser Glu Leu Thr Thr Thr Met Gly Glu Ala Asp
 450 455 460
 1440 gaa gtr gtr aag gtr tgc tgc ctg gtr ctc gtr gtr gtr gtr gtr gtr
 Gly Gly Glu Lys Val Trp Leu Gln Pro Gly Asp Leu Ala Pro Ala Pro
 465 470 475 480
 1488 gtr gac act ccc agc tct gaa gac ttc agt ccc aac tgc agc acc ctc
 Val Asp Thr Pro Ser Ser Glu Asp Phe Arg Pro Asn Cys Ser Thr Leu
 485 490 495
 1536 aac ttc acc tgc gtr gtr tgc tgc aag aac aac ctg gtr gtr gtr gtr
 Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Gln Pro
 500 505 510
 1584 gtr atg ttt ggc cag ctc tgc cag ctg gtr gtr gtr gtr gtr gtr gtr
 Glu Met Phe Ala Gln Leu Ser His Leu Gln Cys Leu Arg Leu Ser His
 515 520 525
 1632 aac tgc atc tgc gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Asn Cys Ile Ser Gln Ala Val Asn Gly Ser Gln Phe Leu Pro Leu Thr
 530 535 540
 1680 gtr ctg gtr gtr cta gac ctg ttc ctc aat aag ctg gtr gtr gtr gtr
 Gly Leu Gln Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His
 545 550 555 560
 1728 gtr gac tca ttc agt gtr cta cga cga gtr gtr gtr gtr gtr gtr gtr
 Glu His Ser Phe Thr Glu Leu Pro Arg Leu Glu Ala Leu Asp Leu Ser
 565 570 575
 1776 ttc aac agc cag ccc ttt gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Tyr Asn Ser Gln Pro Phe Gly Met Gln Gly Val Gly His Asn Phe Ser
 580 585 590
 1824 ttc gtr gtr ctc ctc gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn
 600 605 610
 1872 aac atc ctc agc cca gtr ttc ctc ctc gtr gtr gtr gtr gtr gtr gtr
 Asn Ile His Ser Gln Val Ser Gln Gln Leu Cys Ser Thr Ser Leu Arg
 615 620 625
 1920 gtr ctg gtr ttc gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Ala Leu Asp Phe Ser Gly Asn Ala Leu Gly His Met Trp Ala Glu Gly
 625 630 635 640
 1968 gtr ctc tat ctg ctc ttc ttc cca gtr gtr gtr gtr gtr gtr gtr gtr
 Asp Leu Tyr Leu His Phe Phe Gln Gly Leu Ser Gly Leu Ile Trp Leu
 645 650 655
 2016 gtr tgc ttc cag aac cgc ctg ctc acc ctc ctc ctc ctc ctc ctc ctc
 Asp Leu Ser Gln Asn Arg Leu His Thr Leu Leu Pro Gln Thr Leu Arg
 660 665 670
 2064 aac ctc ccc aag agc cta cag gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Asn Leu Pro Lys Ser Leu Gln Val Leu Arg Leu Arg Asp Asn Tyr Leu
 675 680 685
 2112 gtr ttc ttt aag tgc gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Ala Phe Phe Lys Trp Trp Ser Leu His Phe Leu Pro Lys Leu Glu Val

30

690 ctc gac cty gca gga aac cag cty aag gcc cty acc aac gcc cty 2160
Leu Asp Leu Ala Gly Asn Gln Leu Lys Ala Leu Thr Asn Gly Ser Leu
705 720 725
5 cct gcc gcc acc cgg ctc cgg aag cty gat gcc acc tpc aac gcc atc 2208
Pro Ala Gly Thr Arg Leu Arg Arg Leu Asp Val Ser Cys Asn Ser Ile
725 730 735
acc ttc gty gcc ccc gcc ttc ttc acc aag gcc aag gac cty cga ggg 2256
Ser Phe Val Ala Pro Gly Phe Ser Lys Ala Lys Gln Leu Arg Gln
740 745 750
ctc aac ctt gcc gcc aac gcc ccc aag aca gty gac cac ttc tgg ttt 2304
Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe
755 760 765
ggg ccc cty ggg aat gcc cty caa ata cta gat gta gcc gcc aac cct 2352
15 Gly Pro Leu Ala Ser Ala Leu Gln Ile Leu Asp Val Ser Ala Asn Pro
770 775 780
ctc cac tpc gcc tct tgg ggg gcc ttt atg gcc ttc cty cty ggg gty 2400
Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Gln Val
785 790 795
20 cag gcc gcc gty gcc gcc cty ccc gcc cgg gty aag tct gcc gcc cty 2448
Gln Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro
805 810 815
gcc cag ctc cag gcc ctc gcc atc ttt gca cag gcc cty gcc ctc tpc 2496
Gly Gln Leu Gln Gly Leu Ser Ile Phe Ala Gln Asp Leu Arg Leu Cys
820 825 830
ctg gat ggg gcc ctc ttc tgg gcc tct ttc gcc ctc ttc cty cty gcc 2544
Leu Asp Gln Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu Ala
835 840 845
gty gcc cty gcc cty gcc gty gcc atg cty cat cac ctc tct gcc tgg 2592
30 Val Ala Leu Gly Leu Gly Val Pro Met Leu His Lys Leu Cys Gly Trp
850 855 860
gcc ctc tgg tac tpc ttc cac cty gcc cty gcc tgg ctt ccc tgg cgg 2640
Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg
865 870 875
35 ggg cgg caa agt ggg cga gat ggg gcc gcc cty ccc tac gcc ttc 2688
Gly Arg Gln Ser Gly Arg Asp Gln Asp Ala Leu Pro Tyr Asp Ala Phe
885 890 895
gty gcc ttc gac aac aag cag gcc gca gty gca gcc tgg gga tac aac 2736
Val Val Phe Asp Lys Thr Gln Ser Ala Val Ala Asp Trp Val Tyr Asn
900 905 910
gag ctt cgg ggg cag cty ggg ggg tpc cgt ggg gcc tgg pca ctc cgc 2784
Gln Leu Arg Gly Gln Leu Gln Cys Arg Gly Arg Trp Ala Leu Arg
915 920 925
ctg tpc cty ggg gaa gcc gcc tgg cty cct gcc aac acc ctc ttt ggg 2832
45 Leu Cys Leu Gln Gly Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Gln
930 935 940
aac cty tgg gcc tgg gcc tct gcc acc gcc aag aag cty ttt gty cty 2880
Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu
945 950 955
50 gcc cac aag gcc cag gcc agt gcc ctc cty gcc gcc gcc ttc cty cty 2928
Ala His Thr Asp Arg Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Asn
965 970 975

31

ggc cag cag gcc cty gty ggg gcc gcc aag gcc gcc gty gty cty gty 2976
Ala Gln Gln Arg Leu Leu Gln Asp Arg Lys Asp Val Val Leu Val
980 985 990
atc cty gcc cct gcc gcc gcc gcc ttc gcc tct gty cgg cty gcc cag 3024
5 Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Val Arg Leu Arg Gln
995 1000 1005
gcc ctc tpc gcc cag agt gcc ctc ctc tgg ccc cac cag ccc agt ggt 3072
Arg Leu Cys Arg Gln Ser Val Leu Leu Trp Pro His Gln Pro Ser Gly
1010 1015 1020
10 cag gcc agt ttc tgg gcc cag cty gcc atg gcc cty acc aag gcc aac 3120
Gln Arg Ser Phe Trp Ala Gln Leu Gly Met Ala Leu Thr Arg Asp Asn
1025 1030 1035
cac cac ttc tat aac cgg aac ttc tpc cag gga ccc acc gcc gaa 3165
15 His His Phe Tyr Asn Arg Asn Phe Cys Gln Gly Pro Thr Ala Gln
1045 1050 1055
<21>- 2
<21>- 1055
<21>- PRT
20 <21>- Homo septans
<40>- 2
Met Pro Met Lys Trp Ser Gly Trp Arg Trp Ser Trp Gly Pro Ala Thr
1 5 10 15
25 His Thr Ala Leu Pro Pro Gln Gly Phe Cys Arg Ser Ala Leu His
20 25 30
Pro Leu Ser Leu Leu Val Gln Ala Ile Met Leu Ala Met Thr Leu Ala
35 40 45
Leu Gly Thr Leu Pro Ala Phe Leu Pro Cys Gln Leu Gln Pro His Gly
50 55 60
30 Leu Val Asn Cys Asn Trp Leu Phe Leu Lys Ser Val Pro His Phe Ser
65 70 75
Met Ala Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn
85 90 95
35 Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu
100 105 110
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Pro Val Gly Leu Ser Pro
115 120 125
Met His Phe Pro Cys His Met Thr Ile Gln Pro Ser Thr Phe Leu Ala
130 135 140
40 Val Pro Thr Leu Gln Leu Asn Leu Ser Tyr Asn Asn Ile Met Thr
145 150 155
Val Pro Ala Leu Pro Lys Ser Leu Ile Ser Leu Ser Leu Ser His Thr
165 170 175
45 Asn Ile Leu Met Leu Asp Ser Ala Ser Leu Ala Gly Leu His Ala Leu
180 185 190
Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Tyr Lys Asn Pro Cys Arg
195 200 205
Gln Ala Leu Gln Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu
210 215 220
50 Thr His Leu Ser Leu Lys Tyr Asn Asn Leu Thr Val Val Pro Arg Asn
225 230 235 240

32

Leu Pro Ser Ser Leu Glu Tyr Leu Leu Leu Ser Tyr Asn Arg Ile Val 245
 Lys Leu Ala Pro Glu Asp Leu Ala Asn Leu Thr Ala Leu Arg Val Leu 250
 5 Asp Val Gly Tyr Asn Cys Arg Cys Asp His Ala Pro Asn Pro Cys 265
 Met Glu Cys Pro Arg His Phe Pro Glu Leu His Pro Asp Thr Phe Ser 270
 His Leu Ser Arg Leu Glu Gly Leu Val Leu Lys Asp Ser Ser Leu Ser 285
 10 Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu 300
 Asp Leu Ser Glu Asn Phe Leu Tyr Lys Cys Ile Thr Lys Thr Lys Ala 315
 15 Phe Glu Gly Leu Thr Glu Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr 330
 Glu Lys Arg Val Ser Phe Ala His Leu Ser Leu Ala Pro Ser Phe Gly 335
 Ser Leu Val Ala Leu Lys Glu Leu Asp Met His Gly Ile Phe Phe Arg 340
 20 Ser Leu Asp Glu Thr Thr Leu Arg Pro Leu Ala Arg Leu Pro Met Leu 350
 Glu Thr Leu Arg Leu Glu Met Asn Phe Ile Asn Glu Ala Glu Leu Gly 365
 25 Ile Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn 380
 Arg Ile Ser Gly Ala Ser Glu Leu Thr Ala Thr Met Gly Glu Ala Asp 395
 30 Gly Gly Glu Lys Val Trp Leu Glu Pro Gly Asp Leu Ala Pro Ala Pro 410
 Val Asp Thr Pro Ser Ser Glu Asp Phe Arg Pro Asn Cys Ser Thr Leu 425
 Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Glu Pro 440
 35 Glu Met Phe Ala Glu Leu Ser His Leu Glu Cys Leu Arg Leu Ser His 455
 Asn Cys Ile Ser Glu Ala Val Asn Gly Ser Glu Phe Leu Pro Leu Thr 470
 Gly Leu Glu Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His 485
 40 Glu His Ser Phe Thr Glu Leu Pro Arg Leu Glu Ala Leu Asp Leu Ser 500
 Tyr Asn Ser Glu Pro Phe Gly Met Glu Gly Val Gly His Asn Phe Ser 515
 45 Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn 530
 Asn Ile His Ser Glu Val Ser Glu Glu Leu Cys Ser Thr Ser Leu Arg 545
 Ala Leu Asp Phe Ser Gly Asn Ala Leu Gly His Met Trp Ala Glu Gly 560
 50 Asp Leu Tyr Leu His Phe Phe Glu Gly Leu Ser Gly Leu Ile Trp Leu 575
 Asp Leu Tyr Leu His Phe Phe Glu Gly Leu Ser Gly Leu Ile Trp Leu 590
 Asp Leu Tyr Leu His Phe Phe Glu Gly Leu Ser Gly Leu Ile Trp Leu 605
 Asp Leu Tyr Leu His Phe Phe Glu Gly Leu Ser Gly Leu Ile Trp Leu 620
 Asp Leu Tyr Leu His Phe Phe Glu Gly Leu Ser Gly Leu Ile Trp Leu 635
 Asp Leu Tyr Leu His Phe Phe Glu Gly Leu Ser Gly Leu Ile Trp Leu 650
 Asp Leu Tyr Leu His Phe Phe Glu Gly Leu Ser Gly Leu Ile Trp Leu 665

33

Asp Leu Ser Glu Asn Arg Leu His Thr Leu Leu Pro Glu Thr Leu Arg 665
 Asn Leu Pro Lys Ser Leu Glu Val Leu Arg Leu Arg Asp Asn Tyr Leu 670
 5 Ala Phe Phe Lys Trp Trp Ser Leu His Phe Leu Pro Lys Leu Glu Val 685
 Leu Asp Leu Ala Gly Asn Glu Leu Lys Ala Leu Thr Asn Gly Ser Leu 690
 705 Pro Ala Gly Thr Arg Leu Arg Arg Leu Asp Val Ser Cys Asn Ser Ile 715
 10 Ser Phe Val Ala Pro Gly Phe Phe Ser Lys Ala Lys Glu Leu Arg Glu 730
 Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe 745
 15 Gly Pro Leu Ala Ser Ala Leu Glu Ile Leu Asp Val Ser Ala Asn Pro 760
 Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Glu Val 775
 20 Glu Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro 790
 Gly Glu Leu Glu Gly Leu Ser Ile Phe Ala Glu Asp Leu Arg Leu Cys 805
 Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu Ala 820
 25 Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly Trp 835
 Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg 850
 30 Gly Arg Glu Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala Phe 865
 Val Val Phe Asp Lys Thr Glu Ser Ala Val Ala Asp Trp Val Tyr Asn 880
 Glu Leu Arg Gly Glu Leu Glu Cys Arg Gly Arg Trp Ala Leu Arg 895
 35 Leu Cys Leu Glu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Glu 910
 Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu 925
 40 Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu 940
 Ala Glu Glu Arg Leu Leu Glu Asp Arg Lys Asp Val Val Val Leu Val 955
 11e Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Val Arg Leu Arg Glu 970
 45 Arg Leu Cys Arg Glu Ser Val Leu Leu Trp Pro His Glu Pro Ser Gly 985
 Glu Arg Ser Phe Trp Ala Glu Leu Gly Met Ala Leu Thr Arg Asp Asn 1000
 50 His His Phe Tyr Asn Arg Phe Cys Glu Gly Pro Thr Ala Glu 1015
 His His Phe Tyr Asn Arg Phe Cys Glu Gly Pro Thr Ala Glu 1030
 His His Phe Tyr Asn Arg Phe Cys Glu Gly Pro Thr Ala Glu 1045
 His His Phe Tyr Asn Arg Phe Cys Glu Gly Pro Thr Ala Glu 1060
 His His Phe Tyr Asn Arg Phe Cys Glu Gly Pro Thr Ala Glu 1075

34

Seq.id.no.3
atg ggt ttc tgc cgc agc gcc ctg cac
met gty phe cys arg ser ala leu his
5
ccg ctg tct ctc ctg ctg cag gcc atc atg ctg gcc atg acc ctg gcc 75
Pro Leu Ser Leu Leu Val Gln Ala Ile Met Leu Ala Met Thr Leu Ala
10
ctg ggt acc ttg cct gcc ttc cta ccc tgt gag ctg cag ccc cac gcc 123
Leu Gly Thr Leu Pro Ala Phe Leu Pro Cys Gly Leu Gln Pro His Gly
15
ctg ggt aac tgc aac tgg ctg ttc ctg aag tct gtt gcc cac ttc ttc 171
Leu Val Asn Cys Asn Trp Leu Phe Leu Lys Ser Val Pro His Phe Ser
20
atg gca gca ccc cgt gcc aat gtc acc agc ctt tcc ttg tcc tcc aac 219
Met Ala Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn
25
cgc atc cac cgc cat gat tct gac ctt gcc ccc cgc ctg ccc agc ctg 267
Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu
30
cgg cat ctc aac ctg aag tgg aac tgc cgg cgg gtt gcc ctg acc ccc 315
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Pro Val Gly Leu Ser Pro
35
atg cac ttc ccc tgc cac atg acc atc gag ccc agc acc ttc ttg gct 363
Met His Phe Pro Cys His Met Thr Ile Gly Pro Ser Thr Phe Leu Ala
40
gtg ccc acc ctg gaa gag cta aac ctg agc tac aac aac atc atg act 411
Val Pro Thr Leu Gly Leu Leu Asn Leu Ser Tyr Asn Asn Ile Met Thr
45
gtg cct ggg ctg ccc aaa tcc ctc tta tcc ctg tcc ctg acc cat acc 459
Val Pro Ala Leu Pro Lys Ser Leu Ile Ser Leu Ser Leu Ser His Thr
50
aac atc ctg atg cta gac tct gcc agc ctg gcc ctg cct gcc ctg 507
Asn Ile Leu Met Leu Asp Ser Ala Ser Leu Ala Gly Leu His Ala Leu
55
cgc ttc cta ttc atg gac gcc aac tgt tat tac aag aac ccc tgc agg 555
Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Tyr Lys Asn Pro Cys Arg
60
cag gca ctg gag gtt gcc cgg ggt gcc ctc ctt gcc ctg ggc aac ctc 603
Gln Ala Leu Gly Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu
65
acc cac ctg tca ctc aag tac aac aac ctc act gtt gtt gcc cgg aac 651
Thr His Leu Ser Leu Lys Tyr Asn Asn Leu Thr Val Val Pro Arg Asn
70
ctg cct ttc agc ctg ggg tat ctg ctg ttg tcc tac aac gcc atc gtc 699
Leu Pro Ser Ser Leu Gly Tyr Leu Leu Leu Ser Tyr Asn Arg Ile Val
75
aaa ctg ggt cct gag gac ctg gcc aat ctg acc gcc ctg ggt ctg 747
Lys Leu Ala Pro Gly Asp Leu Ala Asn Leu Thr Ala Leu Arg Val Leu
80
ggt gtt gcc gga aat tgc cgc cgc tgc gcc cac gct ccc aac ccc tgc 795
Asp Val Gly Gly Asn Cys Arg Arg Cys Asp His Ala Pro Asn Pro Cys
250 255 260 265

BIOCOMB-PRO-0155386-1,2

35

atg gag tgc cct cgt cac ttc ccc cag cta cat ccc ggt aac ttc agc 843
Met Gly Cys Pro Arg His Phe Pro Gln Leu His Pro Asp Thr Phe Ser
891
cac ctg agc cgt cct gaa gcc ctg ttg ttg aag agt tct ctc tcc 891
His Leu Ser Arg Leu Gly Leu Val Leu Lys Asp Ser Ser Leu Ser
909
tgg ctg aat gcc agt tgg ttc cgt ggg ctg gga aac ctc cga gtt ctg 909
Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu
957
gac ctg agt gag aac ttc ctc tac aaa tgc atc act aaa acc aag gcc 967
Asp Leu Ser Gly Asn Phe Leu Tyr Lys Cys Ile Thr Lys Thr Lys Ala
1035
ttc cag gcc cta acc cag ctg cgc aag ctt aac ctg ttc ttc aat tac 1035
Phe Gln Gly Leu Thr Gln Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr
1083
caa aag agg ttg ttc ttc gcc cac ctg tct ctg gcc cct tgc ttc ggg 1083
Gln Lys Arg Val Ser Phe Ala His Leu Ser Leu Ala Pro Ser Phe Gly
1131
agc ctg gtc gcc ctg aag gag ctg gac atg cac gcc atc ttc ttc cgc 1131
Ser Leu Val Ala Leu Lys Gly Leu Asp Met His Gly Ile Phe Phe Arg
1179
tea ctc gat gag aac aag ctc cgg cca ctg gcc cgc ctg ccc atg ctc 1179
Ser Leu Asp Gly Thr Thr Leu Arg Pro Leu Ala Arg Leu Met Leu
1227
cag act ctg ggt ctg cag atg aac ttc atc aac aag gcc cag ctg gcc 1227
Gln Thr Leu Arg Leu Gln Met Asn Phe Ile Asn Gln Ala Gln Leu Gly
1275
atc ttc agg gcc ttc cct gcc ctg cgc tac gtt gac ctg tgc gac aac 1275
Ile Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn
1323
cgc atc agc gga gtt tgc gag ctg aca gcc acc atc ggg gag gca gtt 1323
Arg Ile Ser Gly Ala Ser Gly Leu Thr Ala Thr Met Gly Gly Ala Asp
1371
gga ggg gag aag gtc tgg ctg cag cct ggg gac ctt gct cgg gcc cca 1371
Gly Gly Leu Lys Val Trp Leu Gln Pro Gly Asp Leu Ala Pro Ala Pro
1419
gtg gac act ccc agc tct gaa gac ttc agc ccc aac tgc agc acc ctc 1419
Val Asp Thr Pro Ser Ser Gly Asp Phe Arg Pro Asn Cys Ser Thr Leu
1467
aac ttc acc ttg gtt ctg tca cgg aac aac ctg atc acc gtt cag cgg 1467
Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Gln Pro
1515
gag atg ttc gcc cag ctc tgc cac ctg cag tgc ctg gcc ctg agc cac 1515
Glu Met Phe Ala Gln Leu Ser His Leu Gln Cys Leu Arg Ser His
1563
aac tgc atc tgc cag gtc aat ggc tcc cag ttc ctg cgg ctg acc 1563
Asn Cys Ile Ser Gln Ala Val Asn Gly Ser Gln Phe Leu Leu Thr
1611
ggt ctg cag gtt cta gac ctg tcc cac aat aag ctg gac ctc tac cac 1611
Gly Leu Gln Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His
1659
gag cac tca ttc atg gag cta cca cga ctg gag gcc ctg gac ctc agc 1659

BIOCOMB-PRO-0155386-1,2

37

810 gct cct cgt ggc cgt ggc ccc atg cct ctc tgc ggc tgg 820 825
Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly Trp 2523
830 835
gac ctc tgg tac tgc ttc ctc cgt ggc tgg ctc ctc tgg cgg 2571
Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg
845 850
ggg cgg caa agt ggg cga gat gac cct cgt ccc tgc gct ggc ttc
Gly Arg Gln Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala Phe 2619
860 865
gtg gtc ttc gac aac acg cag agc gca gtc gca gac tgg gtc tac aac 2667
Val Val Phe Asp Lys Thr Gln Ser Ala Val Ala Asp Trp Val Tyr Asn
875 880
gag ctt cgg ggg cgg cgt ggc ggc tgc cgt ggg cgc tgg gca ctc cgc 2715
Glu Leu Arg Gly Gln Leu Glu Cys Arg Gly Arg Trp Ala Leu Arg
890 900
ctg tgc cgt ggg gaa cgc cgc tgc cgt cct ggc aac acc ctc ttt gag 2763
Leu Cys Leu Glu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Glu
910 915
aac cgt tgg ggc cca gtc tat ggc agc agc agc cgt ttt gtc cgt 2811
Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu
925 930
gac cac acg cag cgg gtc agt gct ctc ttt ggc ggc ctc cgt cgt 2859
Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu
940 945
gac cag cag cgt cgt cgt ggc agc agc gtc gtc gtc gtc cgt cgt 2907
Ala Gln Gln Arg Leu Leu Glu Asp Arg Lys Asp Val Val Leu Val
955 960
atc cgt agc cct gac ggc cgc tcc cgc tat gtc cgt cgt cgt cgt 2955
11a Leu Ser Pro Asp Gly Arg Ser Arg Tyr Val Arg Leu Arg Gln
970 975
gac ctc tgc cgc cag agt gtc ctc ctc tgg ccc cag cag ccc agt gct 3003
Arg Leu Cys Arg Gln Ser Val Leu Leu Trp Pro His Gln Pro Ser Gly
990 995
cag cgc agc ttc tgg ggc cag cgt ggc atg ggc cgt acc agc gac aac 3051
Gln Arg Ser Phe Trp Ala Gln Leu Gly Met Ala Leu Thr Arg Asp Asn
1005 1010
cac cac ttc tat aac cgg aac ttc tgc cag gga ccc agc ggc gaa 3096
His His Phe Tyr Asn Arg Asn Phe Cys Gln Gly Pro Thr Ala Glu
1020 1025 1030

45 SeqId.No.4

Met Gly Phe Cys Arg Ser Ala Leu His

Pro Leu Ser Leu Leu Val Gln Ala 11a Met Leu Ala Met Thr Leu Ala
10 20 25
Leu Gly Thr Leu Pro Ala Phe Leu Pro Cys Glu Leu Gln Pro His Gly
30 35 40
Leu Val Asn Cys Asn Trp Leu Phe Leu Lys Ser Val Pro His Phe Ser

36

540 545
Glu His Ser Phe Thr Glu Leu Pro Arg Leu Glu Ala Leu Asp Leu Ser
550 555
tac aac agc cag ccc ttt ggc atg cag ggc ggc cac aac ttc agc
Tyr Asn Ser Gln Pro Phe Gly Met Gln Gly Val Gly His Asn Phe Ser
560 565
ttc gtc gct cag cgt acc cgt cgc ctc agc cgt cgt cgc aac aac
Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn
570 575
aac atc cag agc caa gtc tcc cgg cag ctc tgc agt agc tgc cgt cgg
Asn 11a His Ser Gln Val Ser Gln Gln Leu Cys Ser Thr Ser Leu Arg
580 585
gac cgt gac ttc agc ggc aac gca cgt ggc cat atg tgg ggc ggg gga
Ala Leu Asp Phe Ser Gly Asn Ala Leu Gly His Met Trp Ala Glu Gly
590 595
gac ctc tat cgt cag ctc ttc caa ggc cgt agc gct tgc atc tgg cgt
Asp Leu Tyr Leu His Phe Phe Gln Gly Leu Ser Gly Leu Ile Trp Leu
600 605
gac ttt tcc cag aac cgc cgt cgt ccc caa acc cgt cgc
Asp Leu Ser Gln Asn Arg Leu His Thr Leu Leu Pro Gln Thr Leu Arg
610 615
aac ctc ccc aag agc cta cag gtc cgt cgt cgt gac aat tac cgt
Asn Leu Pro Lys Ser Leu Gln Val Leu Arg Leu Arg Asp Asn Tyr Leu
620 625
gac ttt ttt aag tgg tgg agc ctc ctc ttc cgt ccc aac cgt gaa gtc
Ala Phe Phe Lys Trp Trp Ser Leu His Phe Leu Pro Lys Leu Glu Val
630 635
ctc gac cgt gca gga aac cag cgt aag ggc cgt acc aat ggc agc cgt
Leu Asp Leu Ala Gly Asn Gln Leu Lys Ala Leu Thr Asn Gly Ser Leu
640 645
cct gct ggc acc cgt cgt cgt cgt gct gac agc tgc aac agc atc
Pro Ala Gly Thr Arg Leu Arg Arg Leu Asp Val Ser Cys Asn Ser Ile
650 655
agc ttc gtc ggc ccc ggc ttc ttc aag ggc agc ggc cgt cga ggc
Ser Phe Val Ala Pro Gly Phe Phe Ser Lys Lys Lys Glu Leu Arg Glu
660 665
ctc aac ctt agc ggc aac ggc ctc aag aca gtc gac cag ttc tgg ttt
Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe
670 675
ggg ccc cgt ggc agt gct cgt cgt cca ata cta gac gta agc ggc aac cct
Gly Pro Leu Ala Ser Ala Leu Gln 11a Leu Asp Val Ser Ala Asn Pro
680 685
cgt cag tgc gct tgc ggc ggc ttt atg gac ttc cgt cgt ggc gtc
Leu His Cys Ala Gly Ala Phe Met Asp Phe Leu Leu Glu Val
690 695
cag gct ggc cgt gct cgt cgt cgt cgt ggc agt ggc agt ggc cgt
Gln Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro
700 705
gac cag ctc cag ggc ctc agc atc ttc gca cag ggc cgt cgt ctc tgc
Gly Gln Leu Gln Ser Ile Phe Ala Gln Asp Leu Arg Leu Cys
710 715
cgt gct gac gac ctc tcc tgc gac tgc ttc ggc ctc tgc cgt gct
Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu Ala

38

45
Met Ala Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn
60 65 70
Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu
75 80 85
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Pro Val Gly Leu Ser Pro
90 95 100
Met His Phe Pro Cys His Met Thr Ile Glu Pro Ser Thr Phe Leu Ala
110 115 120
Val Pro Thr Leu Glu Glu Leu Asn Leu Ser Tyr Asn Asn Ile Met Thr
125 130 135
Val Pro Ala Leu Pro Lys Ser Leu Ile Ser Leu Ser Leu Ser His Thr
140 145 150
Asn Ile Leu Met Leu Asp Ser Ala Ser Leu Ala Gly Leu His Ala Leu
155 160 165
Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Tyr Lys Asn Pro Cys Arg
170 175 180
Gln Ala Leu Glu Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu
190 195 200
Thr His Leu Ser Leu Lys Tyr Asn Asn Leu Thr Val Val Pro Arg Asn
210 215
Leu Pro Ser Ser Leu Glu Tyr Leu Leu Leu Ser Tyr Asn Arg Ile Val
220 225 230
Lys Leu Ala Pro Glu Asp Leu Ala Asn Leu Thr Ala Leu Arg Val Leu
235 240 245
Asp Val Gly Gly Asn Cys Arg Arg Cys Asp His Ala Pro Asn Pro Cys
250 255 260
Met Glu Cys Pro Arg His Phe Pro Gln Leu His Pro Asp Thr Phe Ser
270 275 280
His Leu Ser Arg Leu Glu Gly Leu Val Leu Lys Asp Ser Leu Ser
285 290 295
Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu
300 305 310
Asp Leu Ser Glu Asn Phe Leu Tyr Lys Cys Ile Thr Lys Thr Lys Ala
315 320 325
Phe Gln Gly Leu Thr Gln Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr
330 335 340
Gln Lys Arg Val Ser Phe Ala His Leu Ser Leu Ala Pro Ser Phe Gly
350 355 360
Ser Leu Val Ala Leu Lys Glu Leu Asp Met His Gly Ile Phe Phe Arg
365 370 375
Ser Leu Asp Glu Thr Thr Leu Arg Pro Leu Ala Arg Leu Pro Met Leu
380 385 390
Gln Thr Leu Arg Leu Gln Met Asn Phe Ile Asn Gln Ala Gln Leu Gly
395 400 405
Ile Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn
410 415 420
Arg Ile Ser Gly Ala Ser Glu Leu Thr Ala Thr Met Gly Glu Ala Asp
430 435 440
Gly Gly Glu Lys Val Trp Leu Gln Pro Gly Asp Leu Ala Pro Ala Pro
445 450 455
Val Asp Thr Pro Ser Ser Glu Asp Phe Arg Pro Asn Cys Ser Thr Leu

39

460 465 470
Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Gln Pro
475 480 485
Glu Met Phe Ala Gln Leu Ser His Leu Gln Cys Leu Arg Leu Ser His
490 495 500
Asn Cys Ile Ser Gln Ala Val Asn Gly Ser Gln Phe Leu Pro Leu Thr
510 515 520
Gly Leu Gln Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His
525 530 535
Glu His Ser Phe Thr Glu Leu Pro Arg Leu Glu Ala Leu Asp Leu Ser
540 545 550
Tyr Asn Ser Gln Pro Phe Gly Met Gln Gly Val Gly His Asn Phe Ser
555 560 565
Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn
570 575 580
Asn Ile His Ser Gln Val Ser Gln Gln Leu Cys Ser Thr Ser Leu Arg
590 595
Ala Leu Asp Phe Ser Gly Asn Ala Leu Gly His Met Trp Ala Glu Gly
605 610 615
Asp Leu Tyr Leu His Phe Phe Gln Gly Leu Ser Gly Leu Ile Trp Leu
620 625 630
Asp Leu Ser Gln Asn Arg Leu His Thr Leu Leu Pro Gln Thr Leu Arg
635 640 645
Asn Leu Pro Lys Ser Leu Gln Val Leu Arg Leu Arg Asp Asn Tyr Leu
650 655 660
Ala Phe Phe Lys Trp Trp Ser Leu His Phe Leu Pro Lys Leu Glu Val
670 675 680
Leu Asp Leu Ala Gly Asn Gln Leu Lys Ala Leu Thr Asn Gly Ser Leu
685 690 695
Pro Ala Gly Thr Arg Leu Arg Arg Leu Asp Val Ser Cys Asn Ser Ile
700 705 710
Ser Phe Val Ala Pro Gly Phe Phe Ser Lys Ala Lys Glu Leu Arg Glu
715 720 725
Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe
730 735 740
Gly Pro Leu Ala Ser Ala Leu Gln Ile Leu Asp Val Ser Ala Asn Pro
750 755 760
Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Val
765 770 775
Gln Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro
780 785 790
Gly Gln Leu Gln Gly Leu Ser Ile Phe Ala Gln Asp Leu Arg Leu Cys
795 800 805
Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu Ala
810 815 820
Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly Trp
830 835 840
Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg
845 850 855
Gly Arg Gln Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala Phe
860 865 870
Val Val Phe Asp Lys Thr Gln Ser Ala Val Ala Asp Trp Val Tyr Asn

40

875
Glu Leu Arg Gly Gln Leu Glu Cys Arg Gly Arg Trp Ala Leu Arg 885
890
895
Leu Cys Leu Glu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Glu 905
910
915
Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu 920
925
Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu 935
940
945
Ala Gln Gln Arg Leu Leu Glu Asp Arg Lys Asp Val Val Val Leu Val 955
960
965
Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Val Arg Leu Arg Gln 975
980
985
Arg Leu Cys Arg Gln Ser Val Leu Leu Trp Pro His Gln Pro Ser Gly 995
1000
1005
Gln Arg Ser Phe Trp Ala Gln Leu Gly Met Ala Leu Thr Arg Asp Asn 1010
1015
His His Phe Tyr Asn Arg Asn Phe Cys Gln Gly Pro Thr Ala Glu 1020
1025
1030

CLAIMS

1. An isolated Toll-like-receptor polypeptide consisting essentially of
 - (i) the amino acid sequence SEQ ID NO: 2;
 - (ii) a variant thereof which has immunomodulatory activity; or
 - (iii) a fragment of (i) or (ii) which has immunomodulatory activity.
2. A polypeptide according to claim 1 wherein the variant (ii) has at least 70% identity to the amino acid sequence of SEQ ID NO: 2.
3. A polypeptide according to claim 2 which has at least 70% identity to the amino acid sequence of SEQ ID NO: 2 over a region of at least 70% of the full-length sequence provided by SEQ ID No.1 and exhibits toll-like receptor functionality.
4. A polypeptide according to claim 1 or claim 2 wherein the variant (ii) has at least 95% identity to the amino acid sequence of SEQ ID NO: 2.
5. A polypeptide according to claim 2 which has at least 95% identity to the amino acid sequence of SEQ ID NO: 2 over a region of at least 60 contiguous amino acids and exhibits toll-like receptor functionality.
6. A polypeptide according to claim 1 wherein the fragment (iii) is a peptide of up to 150 amino acids in length and exhibits toll-like receptor functionality.
7. A polynucleotide encoding a polypeptide according to any one of claims 1-3.
8. A polynucleotide encoding a Toll-like receptor polypeptide which has immunomodulatory activity, which polynucleotide consists essentially of:
 - (a) the nucleic acid sequence of SEQ ID NO: 1;
 - (b) a sequence complementary thereto;

42

- (c) a sequence which hybridises under stringent conditions to a sequence as defined in (a) or (b);
 (d) a sequence that is degenerate as a result of the genetic code to a sequence as defined in (a), (b) or (c); or
 (e) a sequence having at least 80% identity to a sequence as defined in (a), (b), (c) or (d).

9. A polynucleotide according to claim 7 or claim 8 which is mRNA.

10. A polynucleotide according to claim 7 or claim 8 which is DNA.

11. A polynucleotide according to claim 7 or claim 8 which is cDNA.

12. An isolated toll-like receptor polypeptide which is obtainable by expression *in vivo* or *in vitro* of a polynucleotide according to claim 7 or claim 8.

13. A polypeptide according to claim 12 which has the structural features conserved amongst toll-like receptors.

14. An expression vector comprising a polynucleotide sequence according to any one of claims 7 to 11, which is capable of expressing a polypeptide according to any one of claims 1 to 3 or claim 12.

15. An expression vector according to claim 14 which is a plasmid, phage or viral vector.

16. A host cell comprising an expression vector according to claim 14 or claim 15.

43

17. A polyclonal or monoclonal antibody, or a chimera or fragment thereof, which is specific for a polypeptide according to any one of claims 1 to 3.

18. A method for identification of a compound that modulates Toll-like receptor activity, which method comprises contacting a polypeptide according to any one of claims 1 to 3 or claim 12 with a test substance and monitoring for immunomodulatory activity.

19. A compound which modulates Toll-like receptor activity and which is identifiable by a method according to claim 18.

20. A compound according to claim 19 which is a peptide or polypeptide.

21. A compound according to claim 19 which is an oligonucleotide or polynucleotide.

22. A method of treating a subject having an inflammatory or cardiovascular disorder, systemic infection or autoimmune disease that is responsive to Toll-like receptor modulation, which method comprises administering to said subject an effective amount of a compound according to any one of claims 19 to 21 or an antibody according to claim 17.

23. A method according to claim 22 wherein the disorder is a viral, fungal or bacterial infection, asthma, rhinitis, chronic obstructive pulmonary disease (COPD), emphysema, an inflammatory bowel disease such as ulcerative colitis or Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, Alzheimer's disease, atherosclerosis, Multiple Sclerosis, diabetes or septic shock syndrome associated with systemic infection involving gram positive or gram negative bacteria.

24. A polypeptide comprising a fragment or variant of SEQ ID NO: 2, which is capable of inhibiting the activity of TLR9 having the amino acid sequence of SEQ ID NO: 2 or a functional variant thereof, for use in the treatment of an immune or inflammatory disorder.

5

25. A polypeptide according to any one of claims 1 to 3 or claim 12, a polynucleotide according to claim 7 or claim 8 or a compound according to claim 19 for use as an adjuvant.

10

26. The use of a compound according to claim 19 in the manufacture of a medicament for the treatment of an immune or inflammatory disorder.

15

27. The use of a polypeptide comprising a fragment or variant of SEQ ID NO: 2, which is capable of inhibiting the activity of TLR9 having the amino acid sequence of SEQ ID NO: 2 or a functional variant thereof, in the manufacture of a medicament for the treatment of an immune or inflammatory disorder.

20

28. The use of a polypeptide according to any one of claims 1 to 3 or claim 12, a polynucleotide according to claim 7 or claim 8 or a compound according to claim 19 in the manufacture of an adjuvant or vaccine formulation.

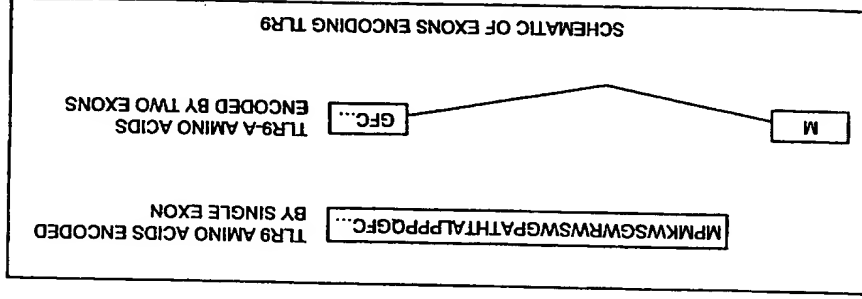
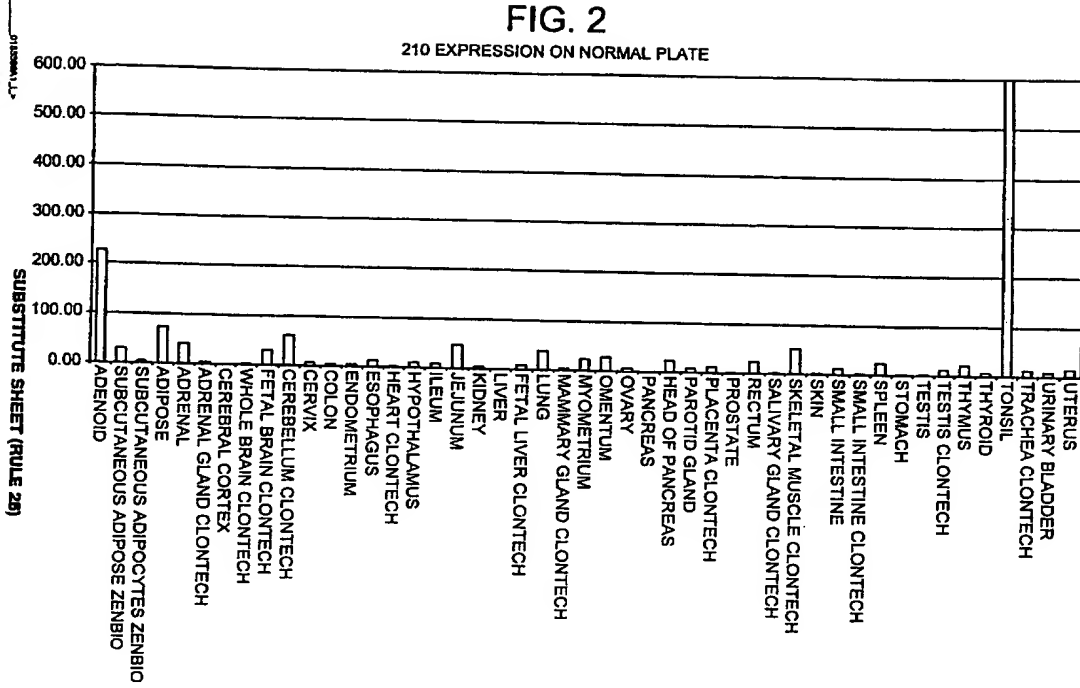


FIG. 1

2 / 5



3 / 5

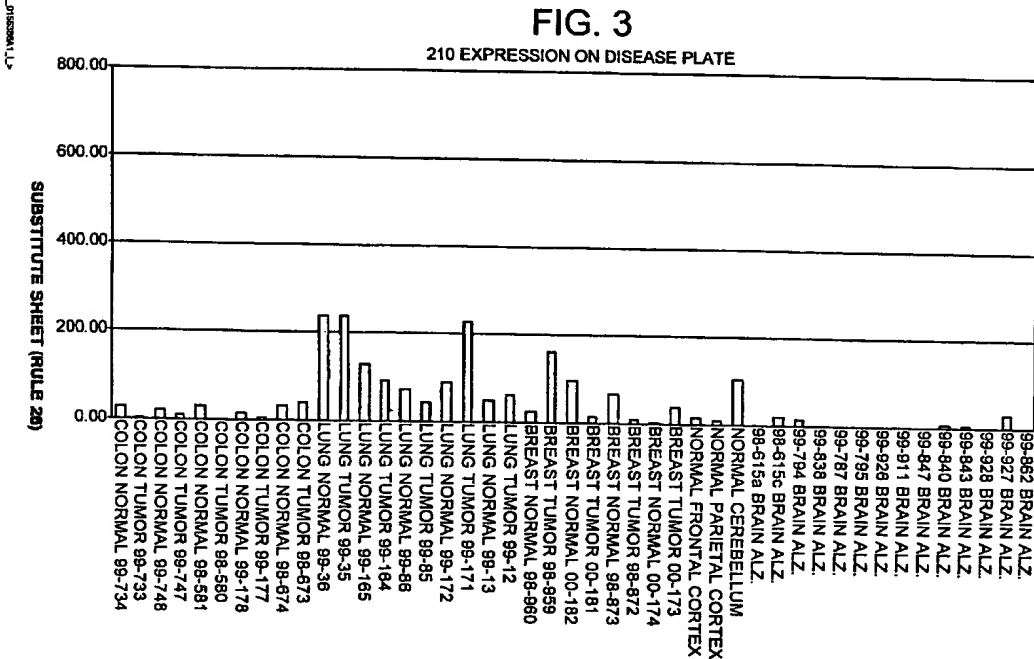


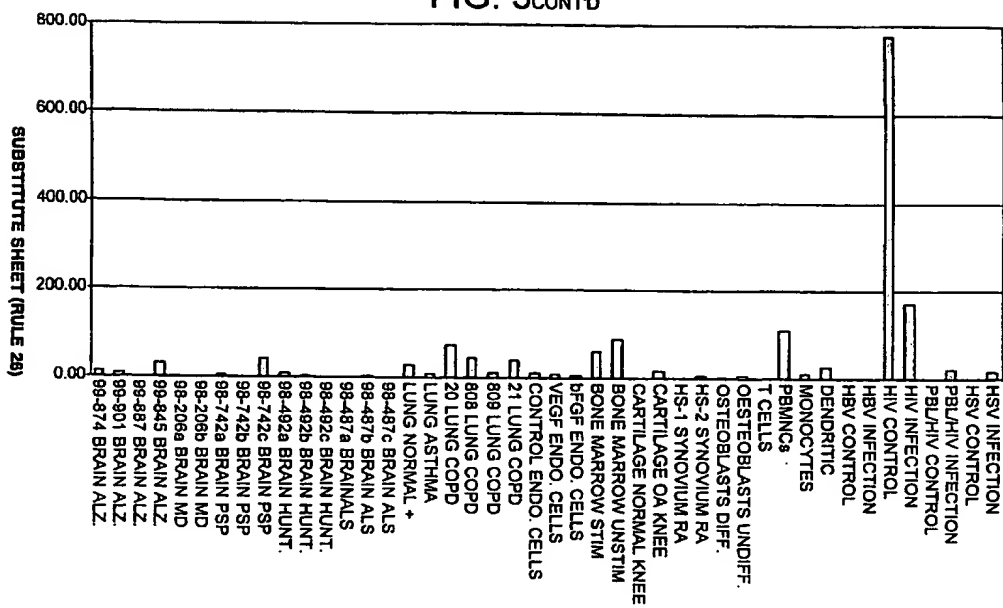
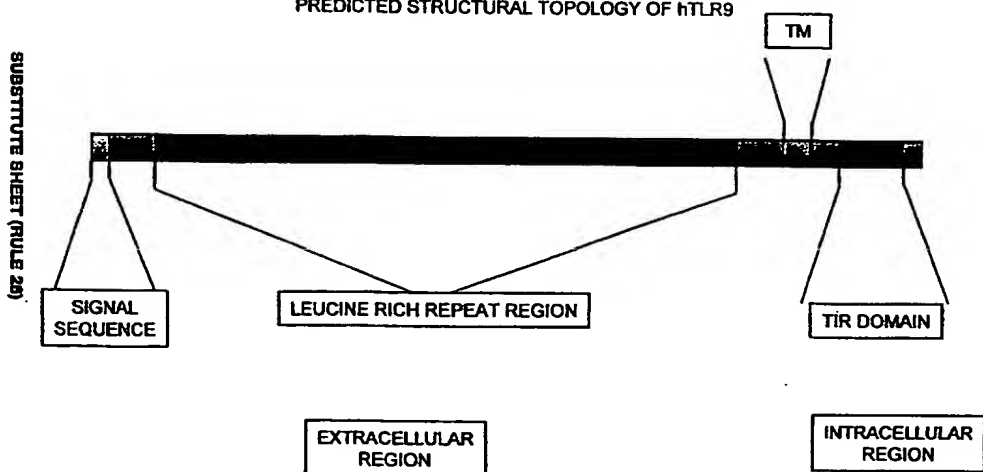
FIG. 3_{CONT'D}

FIG. 4

PREDICTED STRUCTURAL TOPOLOGY OF hTLR9



INTERNATIONAL SEARCH REPORT

Intern. Appl. Application No. PCT/88 01/00299							
<p>A. CLASSIFICATION OF SUBJECT MATTER</p> <p>IPC 7 C12N15/12 C12N15/63 C12N5/10 G01N33/68 C07K14/705 C12N16/28 A61K31/7088 A61K38/17 A61K48/00</p>							
<p>B. FIELD SEARCHED</p> <p>IPC 7 C12N G01N C07K A61K</p>							
<p>C. DOCUMENTS CONSIDERED TO BE RELEVANT</p> <p>Category 1 Citation of document, with indication, where appropriate, of the relevant passages</p> <table border="1"> <thead> <tr> <th>Category</th> <th>Citation of document, with indication, where appropriate, of the relevant passages</th> <th>Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>X</td> <td> <p>NO 98 50547 A (SCHERING CORP) 12 November 1998 (1998-11-12)</p> <p>99.8% identity in 1008 bp overlap between SEQ ID NO 33 of W09850547 and SEQ ID NO 1 99.7% identity in 336 amino acids overlap between SEQ ID NO 34 of W09850547 and SEQ ID NO 2</p> <p>page 56, line 23 -page 60, line 14: claims 1-16; examples 1-11</p> <p>-/-</p> </td> <td>1-18; 22-25, 27, 28</td> </tr> </tbody> </table>		Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	X	<p>NO 98 50547 A (SCHERING CORP) 12 November 1998 (1998-11-12)</p> <p>99.8% identity in 1008 bp overlap between SEQ ID NO 33 of W09850547 and SEQ ID NO 1 99.7% identity in 336 amino acids overlap between SEQ ID NO 34 of W09850547 and SEQ ID NO 2</p> <p>page 56, line 23 -page 60, line 14: claims 1-16; examples 1-11</p> <p>-/-</p>	1-18; 22-25, 27, 28
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.					
X	<p>NO 98 50547 A (SCHERING CORP) 12 November 1998 (1998-11-12)</p> <p>99.8% identity in 1008 bp overlap between SEQ ID NO 33 of W09850547 and SEQ ID NO 1 99.7% identity in 336 amino acids overlap between SEQ ID NO 34 of W09850547 and SEQ ID NO 2</p> <p>page 56, line 23 -page 60, line 14: claims 1-16; examples 1-11</p> <p>-/-</p>	1-18; 22-25, 27, 28					
<p>D. SPECIAL CATEGORIES OF DOCUMENTS:</p> <p>*1* later documents published after the international filing date of the present application and which are cited to substantiate the principle or theory underlying the invention</p> <p>*2* document of particular relevance to the claimed invention</p> <p>*3* document of particular relevance to the claimed invention which is cited to substantiate the publication date of another document</p> <p>*4* document of particular relevance to the claimed invention which is cited to substantiate the publication date of another document</p> <p>*5* document published prior to the international filing date but after the priority date claimed</p>							
<p>E. DATE OF ACTUAL COMPLETION OF THE INTERNATIONAL SEARCH</p> <p>28 June 2001</p>							
<p>F. NAME AND ADDRESS OF THE ISA</p> <p>European Patent Office, P.O. Box 1618, 65001 Luxembourg 2 Tel. (+352-70) 540-5540, Tel. 31 851 850 85 Fax (+352-70) 540-5518</p>							

INTERNATIONAL SEARCH REPORT

Intern. Appl. Application No. PCT/88 01/00299										
<p>C. DOCUMENTS CONSIDERED TO BE RELEVANT</p> <p>Category 1 Citation of document, with indication, where appropriate, of the relevant passages</p> <table border="1"> <thead> <tr> <th>Category</th> <th>Citation of document, with indication, where appropriate, of the relevant passages</th> <th>Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>X</td> <td> <p>DATABASE EMBL 'Online! accession: AC006252 29 December 1998 (1998-12-29) MIZUY D ET AL: "Homo sapiens 3p21.1 contig 9 PAC RPC15-1157M23 (Roswell Park Cancer Institute Human PAC Library) complete sequence." XP002170407</p> <p>100% identity in 3165 bp overlap (between positions 26803-29967) with SEQ ID NO 1 100% identity in 1055 amino acids overlap (between positions 26803-29967) with SEQ ID NO 2 (tfasta)</p> </td> <td>7-11</td> </tr> <tr> <td>P, X</td> <td> <p>HEMMI HIROAKI ET AL: "A Toll-like receptor recognizes bacterial DNA" NATURE, MACMILLAN JOURNALS LTD, LONDON, GB, vol. 408, no. 6813, 7 December 2000 (2000-12-07), pages 740-745, XP002168474 ISSN: 0028-0836 cited in the application the whole document -8 DATABASE EMBL 'Online! accession: AB045180, 13 December 2000 (2000-12-13) AKIRA S ET AL: "Homo sapiens TLR9 mRNA for toll-like receptor 9, complete cds." XP002170408</p> <p>99.7% identity in 3113 bp overlap with SEQ ID NO 1 100% identity in 1031 amino acids overlap with SEQ ID NO 2</p> </td> <td>1-18, 22-25, 27, 28</td> </tr> </tbody> </table>		Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	X	<p>DATABASE EMBL 'Online! accession: AC006252 29 December 1998 (1998-12-29) MIZUY D ET AL: "Homo sapiens 3p21.1 contig 9 PAC RPC15-1157M23 (Roswell Park Cancer Institute Human PAC Library) complete sequence." XP002170407</p> <p>100% identity in 3165 bp overlap (between positions 26803-29967) with SEQ ID NO 1 100% identity in 1055 amino acids overlap (between positions 26803-29967) with SEQ ID NO 2 (tfasta)</p>	7-11	P, X	<p>HEMMI HIROAKI ET AL: "A Toll-like receptor recognizes bacterial DNA" NATURE, MACMILLAN JOURNALS LTD, LONDON, GB, vol. 408, no. 6813, 7 December 2000 (2000-12-07), pages 740-745, XP002168474 ISSN: 0028-0836 cited in the application the whole document -8 DATABASE EMBL 'Online! accession: AB045180, 13 December 2000 (2000-12-13) AKIRA S ET AL: "Homo sapiens TLR9 mRNA for toll-like receptor 9, complete cds." XP002170408</p> <p>99.7% identity in 3113 bp overlap with SEQ ID NO 1 100% identity in 1031 amino acids overlap with SEQ ID NO 2</p>	1-18, 22-25, 27, 28
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.								
X	<p>DATABASE EMBL 'Online! accession: AC006252 29 December 1998 (1998-12-29) MIZUY D ET AL: "Homo sapiens 3p21.1 contig 9 PAC RPC15-1157M23 (Roswell Park Cancer Institute Human PAC Library) complete sequence." XP002170407</p> <p>100% identity in 3165 bp overlap (between positions 26803-29967) with SEQ ID NO 1 100% identity in 1055 amino acids overlap (between positions 26803-29967) with SEQ ID NO 2 (tfasta)</p>	7-11								
P, X	<p>HEMMI HIROAKI ET AL: "A Toll-like receptor recognizes bacterial DNA" NATURE, MACMILLAN JOURNALS LTD, LONDON, GB, vol. 408, no. 6813, 7 December 2000 (2000-12-07), pages 740-745, XP002168474 ISSN: 0028-0836 cited in the application the whole document -8 DATABASE EMBL 'Online! accession: AB045180, 13 December 2000 (2000-12-13) AKIRA S ET AL: "Homo sapiens TLR9 mRNA for toll-like receptor 9, complete cds." XP002170408</p> <p>99.7% identity in 3113 bp overlap with SEQ ID NO 1 100% identity in 1031 amino acids overlap with SEQ ID NO 2</p>	1-18, 22-25, 27, 28								

INTERNATIONAL SEARCH REPORT

Information on patent family members

Intern. Appl. No. PCT/GB 01/00299

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
NO 9850547 A	12-11-1998	AU 7175498 A	27-11-1998
		BR 9808747 A	11-07-2000
		CN 1263555 T	16-08-2000
		EP 0980429 A	23-02-2000
		NO 995458 A	08-11-1999
		PL 336635 A	03-07-2000
		SK 146599 A	11-07-2000
		HU 0001462 A	28-07-2000

From PCT/GB/01/00299 (published 1 July 1999)

INSDOCID: -WFO 018386A1 JJ

International Application No. PCT/GB 01/00299

FURTHER INFORMATION CONTINUED FROM PCT/GB/ 210

Continuation of Box I.2

Claims Nos.: 19-21,26; in part: 22,23,25,28

Claim 19 refers to a compound which modulates Toll-like receptor activity without giving a true technical characterization. Moreover, no such compounds are defined in the application. In consequence, the scope of said claim is ambiguous and vague, and its subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT).

No meaningful search can be carried out for such a purely speculative claim whose wording is, in fact, a mere recitation of the result to be achieved.

The above comment also applies for claims 20,21 and 26; and in part for claims 22,23,25 and 28.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INSDOCID: -WFO 018386A1 JJ

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)